

# What should a statistical mechanics satisfy to reflect nature? \*

Constantino Tsallis<sup>†</sup>

*Centro Brasileiro de Pesquisas Físicas*

*Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro-RJ, Brazil.*

There is no compelling reason imposing that the methods of statistical mechanics should be restricted to the dynamical systems which follow the usual Boltzmann-Gibbs prescriptions. More specifically, ubiquitous natural and artificial systems exhibit complex dynamics, for instance, generic stationary states which are *not* ergodic nor close to it, in any geometrically simple subset of the *a priori* allowed phase space, in any (even extended) trivial sense. A vast class of such systems appears, nevertheless, to be tractable within thermostistical methods completely analogous to the usual ones. The question posed in the title arises then naturally. Some answer to this complex question is advanced in the present review of nonextensive statistical mechanics and its recent connections.

A basic answer to the question of the title could of course be “to make predictions that are experimentally confirmed”. Still, what we would really like to know is what *a priori* mathematical properties or features should a formalism satisfy in order to enable useful statistical predictions in a manner similar to that of Boltzmann-Gibbs (BG) statistical mechanics. This task is either impossible or of extremely high difficulty. In order to be able to formulate an answer which could at least be partially satisfactory, let us analyze the basic structure of the BG theory. Indeed, no logical-deductive path ever existed for proposing a new physical theory, or for generalizing a pre-existing one. In fact, such proposal frequently — perhaps always — occurs on a metaphorical basis. Therefore, the analysis of the structure of the BG theory will hopefully provide us a *metaphor* for formulating a statistical mechanics which could perhaps be more powerful than the one we already have thanks to the genius of Boltzmann, Gibbs, and others. It is perhaps worthy at this stage to explicitly declare that we are talking of a *generalization* of the BG theory, *by no means of an alternative to it*.

## I. LEARNING FROM BOLTZMANN-GIBBS STATISTICAL MECHANICS

### A. Differential equation

Which is the simplest ordinary differential equation? No doubt it is

$$\frac{dy}{dx} = 0 , \quad (1)$$

whose solution (with  $y(0) = 1$ ) is  $y = 1$ . What could be considered as the second in simplicity? Surely

$$\frac{dy}{dx} = 1 , \quad (2)$$

whose solution is  $y = 1 + x$ . And the next one? It is tempting to say

$$\frac{dy}{dx} = y , \quad (3)$$

whose solution is  $y = e^x$ . Its inverse is  $y = \ln x$ , which coincides, by the way, with the celebrated Boltzmann formula (called *Boltzmann principle* by Einstein)

$$S_{BG} = k \ln W , \quad (4)$$

where  $k$  is Boltzmann constant, and  $W$  is the measure of the space where the system is allowed to “live”, taking into account total energy and similar constraints. So, typically, if we have an isolated  $N$ -body Hamiltonian system (microcanonical ensemble in Gibbs notation),  $W$  is the dimensionless Euclidean *measure* (i.e., (hyper)volume) of the fixed-energy Riemann (hyper)surface in phase space (Gibbs’  $\Gamma$ -space) if the system microscopically follows *classical dynamics*, and it is the *dimension* of the associated Hilbert space if the system microscopically follows *quantum dynamics*. In what follows we indistinctively refer to classical or quantum systems. We shall nevertheless use, for simplicity, the wording “phase space” although we shall write down formulas where  $W$  is a natural number.

If we introduce a natural scaling for  $x$  (i.e., if  $x$  carries physical dimensions) we must consider, instead of Eq. (3),

$$\frac{dy}{dx} = ay , \quad (5)$$

\*To appear in a special issue of Physica D entitled *Anomalous Distributions, Nonlinear Dynamics, and Nonextensivity*, eds. H.L. Swinney and C. Tsallis.

<sup>†</sup>tsallis@cbpf.br

in such a way that  $ax$  is a dimensionless variable. The solution is now

$$y = e^{ax} . \quad (6)$$

This differential equation and its solution appear to admit at least three physical interpretations that are crucial in BG statistical mechanics. Let us state them *without proof*. The *first* one is  $(x, y, a) \rightarrow (t, \xi, \lambda)$ , hence

$$\xi = e^{\lambda t} , \quad (7)$$

where  $t$  is time,  $\xi \equiv \lim_{\Delta X(0) \rightarrow 0} \frac{\Delta X(t)}{\Delta X(0)}$  is the *sensitivity to initial conditions*, and  $\lambda$  is the (maximal) Lyapunov exponent associated with a typical phase-space variable  $X$  (the dynamically most unstable one, in fact). This sensitivity to initial conditions (with  $\lambda > 0$ ) is of course the cause of the mixing in phase space which will guarantee *ergodicity*, the well known dynamical justification for the entropy (4).

The *second* physical interpretation is given by  $(x, y, a) \rightarrow (t, \Omega, -1/\tau)$ , hence

$$\Omega = e^{-t/\tau} , \quad (8)$$

where  $\Omega \equiv \frac{\mathcal{O}(t) - \mathcal{O}(\infty)}{\mathcal{O}(0) - \mathcal{O}(\infty)}$ , and  $\tau$  is the characteristic time associated with the *relaxation* of a typical macroscopic observable  $\mathcal{O}$  towards its value at the possible stationary state (*thermal equilibrium* for BG statistical mechanics). This relaxation occurs precisely because of the sensitivity to initial conditions, which guarantees strong chaos (essentially Boltzmann's 1872 *molecular chaos hypothesis*). It was apparently Krylov the first to realize [1], over half a century ago, this deep connection. Indeed,  $\tau$  typically scales like  $1/\lambda$ .

The *third* physical interpretation is given by  $(x, y, a) \rightarrow (E_i, Z p_i, -\beta)$ , hence

$$p_i = \frac{e^{-\beta E_i}}{Z} \left( Z \equiv \sum_{j=1}^W e^{-\beta E_j} \right) , \quad (9)$$

where  $E_i$  is the eigenvalue of the  $i$ -th quantum state of the Hamiltonian (with its associated boundary conditions),  $p_i$  is the probability of occurrence of the  $i$ -th state when the system is at its *macroscopic stationary state* in equilibrium with a thermostat whose temperature is  $T \equiv 1/k\beta$  (canonical ensemble in Gibbs notation). It is a remarkable fact that the *exponential* functional form of the distribution which optimizes the BG generic entropy

$$S_{BG} = -k \sum_{i=1}^W p_i \ln p_i , \quad (10)$$

with the constraints

$$\sum_{i=1}^W p_i = 1 , \quad (11)$$

and

$$\sum_{i=1}^W p_i E_i = U \quad (U \equiv \text{internal energy}), \quad (12)$$

*precisely is the inverse functional form of the same entropy under the hypothesis of equal probabilities*, i.e.,  $p_i = 1/W$  ( $\forall i$ ), hence the *logarithmic* Eq. (4). To the best of our knowledge, there is (yet) no clear generic mathematical linking for this fact, but it is nevertheless true. It might seem at first glance a quite bizarre thing to do that of connecting the standard BG exponential weight to the solution of a (linear) differential equation, in contrast with the familiar procedure consisting in extremizing an entropic functional (Eq. (10)) under appropriate constraints (Eqs. (11) and (12)). It might be helpful to remind to those readers who so think that it is precisely through a differential equation that Planck heuristically found the celebrated black-body radiation law in his October 1900 paper [2], considered by many as the beginning of the path that led to quantum mechanics.

Let us conclude the present remarks by saying that, when we stress that Eqs. (7), (8) and (9) naturally co-emerge within BG statistical mechanics, we only refer to the generic (or more typical) situations, *not to all* the situations. It is known, for example, that relaxation occurs through a power-law function of time at any typical second-order phase transition, whereas the Boltzmann-Gibbs weight remains exponential.

## B. Mean value

The BG entropy (10) can be rewritten as the following mean value:

$$S_{BG} = k \langle \ln \frac{1}{p_i} \rangle , \quad (13)$$

where  $\langle \dots \rangle \equiv \sum_{i=1}^W p_i (\dots)$ . The quantity  $\ln(1/p_i)$  is some times called *surprise* [3] or *unexpectedness* [4]. We notice that the averaged quantity has the *same* functional form as that corresponding to the equal probability case Eq. (4), where  $1/p_i$  plays the role of  $W$ .

## C. Entropy composition law for independent systems

Let us consider systems  $A$  and  $B$  as probabilistically independent, i.e., such that  $p_{ij}^{A+B} = p_i^A p_j^B$  ( $\forall (i, j)$ ). We can immediately prove that entropy (10) satisfies the following property

$$S_{BG}(A+B) = S_{BG}(A) + S_{BG}(B) , \quad (14)$$

referred from now on as *extensivity*. This property is sometimes referred to as *additivity*, reserving the word

*extensivity* for the infinitely many body systems; we will for simplicity not make such distinction here.

The *linear* property (14) of course encompasses the fact that, since  $W_{A+B} = W_A W_B$ , whenever we have equal probabilities, the logarithmic form (4) is absolutely fitted. For example, if we have  $N$  independent coins (or dices), it is  $W = 2^N$  (or  $6^N$ ), hence  $S_{BG} = Nk \ln 2$  (or  $S_{BG} = Nk \ln 6$ ). If we have, as another example, a  $d = 3$  regular lattice with ferromagnetic Heisenberg interactions between first neighbors at very high temperature, it is  $W \sim A\rho^N$  (with  $A > 0$ ,  $\rho > 1$ , and  $N \rightarrow \infty$ ), hence  $S_{BG} \sim Nk \ln \rho$ . In all these cases, we have  $S_{BG} \propto N$ , which precisely fits the Clausius concept of thermodynamic entropy. We shall discuss in Section II what can we do when the ubiquitous (but not obligatory) behavior  $W(N) \sim \rho^N$  (with  $N \gg 1$ ) is drastically violated, e.g., when  $W \propto N^\gamma$ , with  $\gamma > 0$ , which *also* appears to be ubiquitous in both natural and artificial systems.

## D. Concavity

Let us consider two probability distributions  $\{p_i\}$  and  $\{p'_i\}$  for a given system ( $i = 1, 2, \dots, W$ ). We shall define an *intermediate* distribution as follows:

$$p''_i \equiv \mu p_i + (1 - \mu) p'_i \quad (0 < \mu < 1), \quad (15)$$

An entropic functional  $S(\{p_i\})$  is said *concave* if and only if

$$S(\{p''_i\}) \geq \mu S(\{p_i\}) + (1 - \mu) S(\{p'_i\}) \quad (\forall \{p_i\}, \forall \{p'_i\}, \forall \mu). \quad (16)$$

It can be easily shown that  $S_{BG}$  is concave. This property generically yields, within *BG* statistical mechanics, *thermodynamical stability*, i.e., stability of the system with regard to energetic perturbations. For the canonical ensemble this implies  $\partial^2 S_{BG} / \partial U^2 \leq 0$  (i.e., positive specific heat). For the microcanonical ensemble, the situation can in fact be more complex (see, for instance, [5]). Also, it is this property which makes that if we put into thermal contact two systems which are initially at different temperatures (of the same sign, usually positive), thermal equilibrium occurs at a temperature which is necessarily *intermediate* between the two initial ones. And this occurs with a variation of the total entropy which is necessarily an *increase*. We therefore see how central the concavity of the entropy is for both the 0<sup>th</sup> and the 2<sup>nd</sup> principles of thermodynamics.

## E. Stability or experimental robustness

Lesche [6] addressed in 1982 a very basic (and, curiously enough, widely unknown) property. A necessary condition for a (positive) statistical functional  $O(\{p_i\})$  to be a physical quantity is [6] that, under arbitrary small variations of the probabilities  $\{p_i\}$ , its relative variation remains small. We then say that  $O(\{p_i\})$  is *stable* [6] (we

shall hereafter also use the expression *experimentally robust*, in order to avoid confusion with the thermodynamic stability mentioned in the previous subsection). To be more precise, if we consider two probability sets  $\{p_i\}$  and  $\{p'_i\}$  associated with  $W$  microstates, the measure of the size of the deformation can be defined as follows [6]

$$\|p - p'\| = \sum_{i=1}^W |p_i - p'_i|, \quad (17)$$

The condition of stability or experimental robustness of  $O(\{p_i\})$  is then given by

$$\|p - p'\| < \delta_\varepsilon \implies R \equiv \left| \frac{O(\{p_i\}) - O(\{p'_i\})}{O_{max}} \right| < \varepsilon, \quad (18)$$

for any  $\varepsilon > 0$ , with  $\delta_\varepsilon > 0$  being independent from  $W$ ,  $O_{max}$  being the maximal value that  $O(\{p_i\})$  can attain. This implies, in particular, that  $\lim_{\varepsilon \rightarrow 0} \lim_{W \rightarrow \infty} R = 0$ .

It can be shown [6] that  $S_{BG}$  (for which  $O_{max} = k \ln W$ ) is experimentally robust.

## F. Finite entropy production per unit time

Let us consider the allowed phase space (assumed continuous and finite) of our system, and let us partition it in  $W \gg 1$  little cells. Assume next an ensemble of  $M \gg 1$  initial conditions, all inside one of the  $W$  little cells, and let the microscopic dynamics of the system to run. If the system is strongly chaotic (i.e., if it has positive Lyapunov exponents), the points will quickly spread everywhere inside the allowed region. We might define a set of probabilities  $p_i \equiv M_i(t)/M$  ( $i = 1, 2, \dots, W$ ), where  $M_i(t)$  is the number of points inside the  $i$ -th cell at time  $t$  ( $\sum_{i=1}^W M_i(t) = M$ ,  $\forall t$ ). The entropy  $S_{BG}(t)$  will be zero at  $t = 0$ , and will start increasing as  $t$  goes on, finally saturating at some value (which is  $k \ln W$  if the system is *ergodic*; more precisely, if it equally visits all regions of the allowed phase space). The quantity

$$K_{BG} \equiv \lim_{t \rightarrow \infty} \lim_{W \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{S_{BG}(\{p_i(t)\})/k}{t} \quad (19)$$

will in general be *finite* and equal to the Kolmogorov-Sinai entropy (which constitutes a slightly different, more familiar in the mathematical community, definition of the same concept). This finiteness follows from the Pesin theorem, which can be rewritten essentially as follows

$$K_{BG} = \sum_r \lambda_r, \quad (20)$$

where  $r$  runs over all the positive Lyapunov exponents of the system.

## II. USING WHAT WE LEARNT FROM BOLTZMANN-GIBBS STATISTICAL MECHANICS TO GENERALIZE IT

There are several other properties than those discussed above, which also specifically characterize BG statistical mechanics, but we shall restrict the present analysis to only those, i.e., differential equations, mean value, entropy composition law, concavity, experimental robustness, and entropy production. For further — and consistent — characterizations, see [7]. As we already mentioned, there is of course no logical-deductive manner to generalize a physical theory. Or, to put it on more general grounds, there is no generic or unique way of generalizing a logically consistent set of axioms into another one which also is logically consistent and which, by construction, recovers the original one as a particular case. It is therefore only metaphorically that we shall use, in what follows, the mathematical structure of BG statistical mechanics in order to generalize it.

### A. Differential equations

Can we unify Eqs. (1), (2) and (3) in a *single* differential equation? Yes, through

$$\frac{dy}{dx} = a + by . \quad (21)$$

Can we do it *minimally*, with only one parameter? (instead of two, namely  $a$  and  $b$ ). Yes, *out of linearity*, through

$$\frac{dy}{dx} = y^q \quad (q \in \mathcal{R}) \quad (22)$$

Eqs. (1), (2) and (3) are respectively recovered for  $q \rightarrow -\infty$ ,  $q = 0$  and  $q = 1$ . The solution of Eq. (22) (with  $y(0) = 1$ ) is given by

$$y = [1 + (1 - q)x]^{1/(1-q)} \equiv e_q^x \quad (e_1^x = e^x) . \quad (23)$$

The inverse function of the  $q$ -exponential is the  $q$ -logarithm, defined as follows

$$y = \frac{x^{1-q} - 1}{1 - q} \equiv \ln_q x \quad (\ln_1 x = \ln x) . \quad (24)$$

It should be clear that *these arguments about differential equations have by no means a provatory nature nor intention*. They are given here to provide some specific “feeling” about *linearity* and *nonlinearity*, thus providing some intuitive plausibility to the generalization that we propose in what follows. On these grounds, we may expect the Boltzmann principle, Eq. (4), to be generalized, for equal probabilities, as follows

$$S_q(p_i = 1/W, \forall i) = k \ln_q W = k \frac{W^{1-q} - 1}{1 - q} , \quad (25)$$

As for the BG case, if  $x$  carries a physical dimension, we must consider, instead of Eq. (22),

$$\frac{dy}{dx} = a_q y^q \quad (a_1 = a) , \quad (26)$$

hence

$$y = e_q^{a_q x} , \quad (27)$$

As for the BG case, we expect this solution to admit at least three different physical interpretations. The first one corresponds to the sensitivity to initial conditions

$$\xi = e_q^{\lambda_q t} , \quad (28)$$

where  $\lambda_q$  generalizes the Lyapunov exponent or coefficient. Expression (28) was conjectured in 1997 [8], and, for unimodal maps, proved recently [9,10].

The second interpretation corresponds to relaxation, i.e.,

$$\Omega = e_q^{-t/\tau_q} , \quad (29)$$

There is (yet) no proof of this property, but there are several verifications (see, for instance, [11] for a quantum chaotic system).

The third interpretation corresponds to the energy distribution at the stationary state, i.e.,

$$p_i = \frac{e_q^{-\beta_q E_i}}{Z_q} \left( Z_q \equiv \sum_{j=1}^W e_q^{-\beta_q E_j} \right) . \quad (30)$$

This is precisely the form that comes out from the optimization of the generic entropy  $S_q$  under appropriate constraints [12–14]. This form has been observed in a large variety of situations (see [15] for mini-reviews).

Before closing this subsection, let us stress that there is no reason for the values of  $q$  appearing in Eqs. (28), (29) and (30) be the same. Indeed, if we respectively denote them by  $q_{sen}$  (*sen* stands for *sensitivity*),  $q_{rel}$  (*rel* stands for *relaxation*) and  $q_{stat}$  (*stat* stands for *stationary state*), we typically (but not necessarily) have that  $q_{sen} \leq 1$ ,  $q_{rel} \geq 1$  and  $q_{stat} \geq 1$ . The possible connections between all these entropic indices are not (yet) known in general. However, for the edge of chaos of the  $z$ -logistic maps (see [16–19] and references therein) we do know some important properties. If we consider the multifractal  $f(\alpha)$  function, the fractal or Hausdorff dimension  $d_f$  corresponds to the maximal height of  $f(\alpha)$ ; also, we may denote by  $\alpha_{min}$  and  $\alpha_{max}$  the values of  $\alpha$  at which  $f(\alpha)$  vanishes (with  $\alpha_{min} < \alpha_{max}$ ). It has been proved [9,16] that

$$\frac{1}{1 - q_{sen}} = \frac{1}{\alpha_{min}} - \frac{1}{\alpha_{max}} . \quad (31)$$

Moreover, there is some numerical evidence [18] suggesting

$$\frac{1}{q_{rel} - 1} \propto (1 - d_f) . \quad (32)$$

Unfortunately, we know not much about  $q_{stat}$ , but it would not be surprising if it was closely related to  $q_{rel}$ . They could even coincide, in fact.

### B. Mean value

Since we have seen in the previous subsection that the logarithmic function naturally generalizes into the  $q$ -logarithmic one, let us define

$$S_q = k \langle \ln_q \frac{1}{p_i} \rangle , \quad (33)$$

where we may call  $\ln_q(1/p_i)$  the  $q$ -surprise or  $q$ -unexpectedness. Then, using Eq.(24), it is straightforward to obtain

$$S_q = k \frac{1 - \sum_{i=1}^W p_i^q}{q - 1} \quad (S_1 = S_{BG}) , \quad (34)$$

which is the entropy on which we shall base the present generalization of BG statistical mechanics [12–14]. This entropy of course recovers Eq. (25) for equal probabilities.

### C. Entropy composition law for independent systems

If we consider now the same two probabilistically independent systems  $A$  and  $B$  that we assumed before, we straightforwardly obtain

$$\frac{S_q(A+B)}{k} = \frac{S_q(A)}{k} + \frac{S_q(A)}{k} + (1-q) \frac{S_q(A)}{k} \frac{S_q(B)}{k} . \quad (35)$$

We re-obtain Eq. (14) in the limit  $(1-q)/k \rightarrow 0$ . Since  $S_q$  is always nonnegative, we see that, if  $q < 1$  ( $q > 1$ ), we have that  $S_q(A+B) > S_q(A) + S_q(B)$  ( $S_q(A+B) < S_q(A) + S_q(B)$ ), which shall be referred as the *superextensive* (*subextensive*) case. It is from this property that the expression *nonextensive statistical mechanics* was coined.

Let us address now the case  $W \propto N^\gamma$  ( $\gamma > 0$ ;  $N \gg 1$ ) that we mentioned earlier. If we replace this into Eq. (25) we obtain  $S_q \propto N^{\gamma(1-q)}$  if  $q < 1$ . Consequently, it exists a special and unique value of  $q$ , namely  $q^* = 1 - (1/\gamma) < 1$ , for which we have  $S_{q^*} \propto N$ , once again in agreement with the Clausius idea of entropy! This only occurs because the case  $W \propto N^\gamma$  is incompatible with any hypothesis of probabilistic independence (even in the limit  $N \rightarrow \infty$ ). Amusingly enough, we see that it might happen that the entropy  $S_q$  with  $q \neq 1$ , which is *nonextensive* for independent systems, can be *extensive* when we consider a special class of systems which include strong dependence. More explicitly, if we asymptotically have

$W \propto \rho^N$ , it follows (for  $N_A$  and  $N_B$  large and independent)  $W(A+B) \propto \rho^{N_A+N_B} = \rho^{N_A} \rho^{N_B} \propto W(A)W(B)$ , hence we must choose  $q^* = 1$  to have a finite limit for the entropy per particle. Whereas, if we asymptotically have  $W \propto N^\gamma$ , it follows that  $W(A+B) \propto (N_A + N_B)^\gamma \neq N_A^\gamma N_B^\gamma \propto W(A)W(B)$ , and the way for still having a *finite* limit for the entropy per particle is choosing  $q^* = 1 - (1/\gamma)$ .

It is interesting to notice [10] that if we replace  $W \propto N^\gamma$  into  $S_{2-q}$  we obtain  $S_{2-q} \propto N^{\gamma(q-1)}$ , hence  $S_{2-q^*} \propto N$  if  $q^* = 1 + (1/\gamma) > 1$ . Naturally  $q^{**} = 2 - q^*$ .

In fact, this remark can be trivially generalized. Consider  $\kappa(q)$ , where  $\kappa$  is any monotonic continuous function satisfying  $\kappa(1) = 1$ . If we focus on the decreasing ones, an example of such function is the one already considered, namely  $\kappa = 2 - q$ . Another one is  $\kappa = 1/q$ . Other possibilities can be obtained by successively applying the two just mentioned, i.e.,  $\kappa = 1/(2 - 1/q) = q/(2q - 1)$ , or, the other way around,  $\kappa = 2 - 1/(2 - q) = (3 - 2q)/(2 - q)$ . (In fact, the four possibilities we have just considered belong to the class  $\kappa = (a + bq)/[(a + 2b)q - b]$  with  $(a, b) \in \mathcal{R}^2$ , which is symmetric with regard to the axis  $\kappa = q$ ). Then, for the case  $W \propto N^\gamma$  we are considering, we have that  $S_{\kappa(q^*)} \propto N$  with  $q^* = \kappa^{-1}(1 - (1/\gamma))$ ,  $\kappa^{-1}$  denoting the inverse function. If  $\kappa(q)$  is an increasing (decreasing) function, we have that  $q^* < 1$  ( $q^* > 1$ ). The considerations we have done in this paragraph might be not unrelated with the possible connections between the entropic indices  $q_{rel}$ ,  $q_{stat}$  and  $q_{stat}$  mentioned earlier.

Finally, let us remind that Eq. (35) can be rewritten in the following extensive form [12]:

$$S_q^R(A+B) = S_q^R(A) + S_q^R(B) , \quad (36)$$

where

$$S_q^R \equiv \frac{\ln[1 + (1-q)S_q/k]}{1-q} = \frac{\ln \sum_{i=1}^W p_i^q}{1-q} \quad (37)$$

is the Renyi entropy. This entropy is extremal at  $1/W$  (equal probabilities), and attains the value  $\ln W$  for *all* values of  $q$ . Clearly, this fact makes it useless for having a finite entropy per particle in the case  $W \propto N^\gamma$ . It is, nevertheless, an interesting functional for geometrically characterizing multifractals, as long known.

### D. Concavity

It can be shown that  $S_q(\{p_i\})$  is a concave (convex) functional for *all* positive (negative) values of  $q$ . It also follows that, for the canonical ensemble, the specific heat is necessarily positive for  $q \geq 1$ , but not necessarily for  $0 < q < 1$ .

Let us mention that Renyi entropy is concave for  $q \leq 1$ , but has no definite concavity (or convexity) for  $q > 1$ . We consider this as a serious drawback for the use of Renyi entropy with  $q > 1$  for thermostistical and thermodynamical purposes.

## E. Stability or experimental robustness

Abe has recently proved [20] a remarkable fact, namely that  $S_q$  is stable (experimentally robust) for *all* positive values of  $q$ . In contrast, Renyi entropy is unstable (experimentally fragile) for all values of  $q \neq 1$  [6]. Once again, we consider this a further serious drawback for the use of the Renyi functional as a basis for thermodynamics.

## F. Finite entropy production per unit time

If the dynamics of the system is such that the Lyapunov exponents vanish, Eq. (20) provides  $0 = 0$ , which clearly is poorly informative. We would like to “unfold” this trivial equality, and know more about those zeros. The functional  $S_q$  has been shown to enable precisely this in a variety of situations, including various low- and high-dimensional nonlinear dynamical systems, both conservative and dissipative, both discrete and continuous in time and space. The situation can be summarized as follows.

We generalize definition (19) as follows:

$$K_q \equiv \lim_{t \rightarrow \infty} \lim_{W \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{S_q(\{p_i(t)\})/k}{t} \quad (K_1 = K_{BG}) . \quad (38)$$

What has been verified for vast classes of  $K_{BG} = 0$  nonlinear dynamical systems (see, for instance, [21]) is that a special and unique value of  $q$  exists (coincident with  $q_{sen} < 1$  every time checking has been performed) such that  $K_{q_{sen}}$  is finite, whereas  $K_q$  vanishes (diverges) for all values of  $q > q_{sen}$  ( $q < q_{sen}$ ). In other words, for such systems, the entropy whose production per unit time is *finite* is  $S_{q_{sen}}$ , *not*  $S_{BG}$ .

The full characterization of the systems so behaving, and the full comprehension of their dynamical details in what concerns wandering in phase space, are still lacking. The scenario seems nevertheless relatively clear. If a chaotic sea (frequently single connected) exists in phase space, and we perform an average over many initial conditions (i.e., many choices for the little cell, among the  $W$  cells existing in the partition, where we initially place the  $M$  points) all over the entire allowed phase space, then the  $q_{sen} = 1$  entropy production is finite, *independently on whether the measure of occupancy is uniform* (i.e., ergodic in the entire allowed phase space) or not. But if the evolution in phase space is such that the system remains sensibly long times in regions where the structure is multifractal-like (possibly scale-free-like, in the sense of Barabasi et al, see [22] and references therein), then one expects  $q_{sen} < 1$ . More details will be shown later in this paper.

The generalization of the Pesin theorem (Eq. (20)) along the present lines was conjectured in 1997 [8], and has been recently proved by Baldovin and Robledo [10] for unimodal one-dimensional maps. More precisely, they have proved that

$$K_{q_{sen}} = \lambda_{q_{sen}} \quad (39)$$

We believe this to be one of the cornerstones of the entire theory.

It is clear that the comments we did earlier concerning the behavior of  $S_q$  as a function of  $N$  are applicable here as a function of  $t$ . In other words, the entropy production per unit time is finite also for  $S_{\kappa(q)}(t)$ , i.e., if  $K_{q_{sen}}$  is finite, the same occurs for  $K_{\kappa(q_{sen})}$ . Also, it is clear from Eq. (37) that for all those  $q \neq 1$  systems for which  $S_q(t)$  is asymptotically linear, Renyi entropy is *not*. Therefore, for such systems, there is no value of  $q$  (excepting of course whenever  $q_{sen} = 1$ ) for which  $S_q^R$  asymptotically could yield a finite entropy production per unit time.

## III. THE CANONICAL ENSEMBLE AND CONNECTION TO THERMODYNAMICS

In the previous Sections we have addressed the question of *which one*, among the infinite possible generalizations of  $S_{BG}$ , we want to *postulate* as the basis for generalizing BG statistical mechanics. We have decided it will be  $S_q$ . Now, if we have an entropic form, in principle *any* entropic form, and are interested in cybernetics, control theory, information theory, and related matters, there are many things that we can do just with that. If we are, however, physicists, we might naturally think of doing statistical mechanics and thermodynamics. In other words, we will consider the *energy* (which represents the physical support of the system) in addition to the *entropy* (which represents our information about that physical support). To do so, we shall first address what we consider to be the “Sancta Sanctorum” of statistical mechanics: conservative Hamiltonian systems, i.e., the systems Boltzmann and Gibbs themselves had in mind. More specifically, we shall address a large system in contact with an even (much) larger thermostat (*canonical ensemble*). Following along Gibbs’ variational path, we shall extremize  $S_q$  with the norm constraint (11), and with a supplementary constraint related to the Hamiltonian  $\mathcal{H}$  of the system, namely [14]

$$\langle \mathcal{H} \rangle_q \equiv \frac{\sum_{i=1}^W p_i^q E_i}{\sum_{j=1}^W p_j^q} = U_q , \quad (40)$$

with  $\langle \dots \rangle_1 = \langle \dots \rangle$ ;  $P_i \equiv p_i^q / \sum_{j=1}^W p_j^q$  is referred to as the *escort distribution* [23], and  $\{E_i\}$  is the set of eigenvalues associated with  $\mathcal{H}$  and the corresponding boundary conditions. We discuss below the reasons which make desirable the use of  $P_i$  instead of just  $p_i$  for defining the energy constraint. The extremizing distribution corresponds to (meta)equilibrium and is straightforwardly shown to be given by

$$p_i = \frac{e_q^{-\beta_q(E_i - U_q)}}{\bar{Z}_q} , \quad (41)$$

with

$$\bar{Z}_q \equiv \sum_{j=1}^W e_q^{-\beta_q(E_j - U_q)} , \quad (42)$$

and

$$\beta_q \equiv \frac{\beta}{\sum_{j=1}^W p_j^q} , \quad (43)$$

$\beta \equiv 1/kT$  being the Lagrange parameter associated with constraint (40). We easily verify that  $q = 1$  recovers the standard BG weight,  $q > 1$  implies in a *power-law tail* at high values of  $E_i$ , and  $q < 1$  implies in a *cutoff* at high values of  $E_i$ . The (meta)equilibrium distribution (41) can be rewritten as follows:

$$p_i = \frac{e_q^{-\beta'_q E_i}}{Z'_q} , \quad (44)$$

with

$$Z'_q \equiv \sum_{j=1}^W e_q^{-\beta'_q E_j} , \quad (45)$$

and

$$\beta'_q \equiv \frac{\beta_q}{1 + (1 - q)\beta_q U_q} . \quad (46)$$

This form is particularly convenient for many applications where comparison with experimental or computational data is involved.

From the preceding results the connection to thermodynamics can be derived. In fact, the entire Legendre transformation structure of thermodynamics is  $q$ -invariant. In particular, it can be proved that

$$\frac{1}{T} \equiv k\beta = \frac{\partial S_q}{\partial U_q} ,$$

as well as

$$F_q \equiv U_q - \frac{S_q}{k\beta} = -\frac{1}{\beta} \ln_q Z_q ,$$

where

$$\ln_q Z_q = \ln_q \bar{Z}_q - \beta U_q .$$

Also, it can be proved that

$$U_q = -\frac{\partial}{\partial \beta} \ln_q Z_q ,$$

as well as

$$C_q \equiv T \frac{\partial S_q}{\partial T} = \frac{\partial U_q}{\partial T} = -T \frac{\partial^2 F_q}{\partial T^2} .$$

The form adopted for constraint (40) (instead of the usual Eq. (12)) is at first sight astonishing, and surely demands clarification. It satisfies a remarkable set of (intertwined) properties, which we list now.

(i) It satisfies, exactly as it happens when using  $p_i$ , the basic property that the mean value of a constant is the same constant. This was not so in the version developed in 1991 [13]. Chronologically speaking, to satisfy this property only became compelling when it became clear that the  $q$ -generalization concerned a problem of lack of ergodicity, and not any unusual norm-preservation anomaly.

(ii) It makes that the addition of *microscopic* energies of independent systems preserves, at the *macroscopic* level, *exactly the same form*. In other words,  $E_{ij}^{A+B} = E_i^A + E_j^B$  with  $p_{ij}^{A+B} = p_i^A p_j^B$  implies  $U_q^{A+B} = U_q^A + U_q^B$ . This property surely has an important role for the present generalization to satisfy the first principle of thermodynamics. This property was not satisfied in the 1991 version [13].

(iii) The distribution (41) obtained for the stationary state is invariant under change of the zero of energies, i.e., under uniform translation of the microscopic energies. In more specific terms, if we add  $E_0$  to all energies  $E_i$ , the same  $E_0$  is, through Eq. (40), added to  $U_q$ , hence  $E_i - U_q$  does not change, *nor does*  $p_i$ . This property was not explicitly satisfied in the early 1988 and 1991 versions [12,13].

(iv) Abe and Rajagopal showed [24] that the traditional steepest descent method (long ago used by Darwin and Fowler to discuss BG statistics) naturally leads to the escort distribution in the energy constraint. It is basically related to the simple property  $de_q^x/dx = (e_q^x)^q$ .

(v) The form of constraint (40) makes that, under the optimization of  $S_q$ , the Lagrange  $\alpha$ -parameter (the one associated with the norm constraint) appears in a function which naturally *factorizes* out of the sum over all states  $\{i\}$ . This property warrants the definition of a partition function which, as usually, depends on the Lagrange  $\beta$ -parameter but *not on*  $\alpha$ . This interesting mathematical fact was present in the 1991 version, but not in the version developed in 1988.

(vi) If we express the two  $q > 1$  canonical-ensemble constraints in the space of the energy, for high energy they asymptotically behave as indicated now:

$$\begin{aligned} \sum_i p_i &\simeq \int_{constant}^{\infty} dE g(E) p(E) \\ &\propto \int_{constant}^{\infty} dE g(E) / E^{1/(q-1)} \end{aligned}$$

and

$$\begin{aligned} \sum_i p_i^q E_i / \sum_j p_j^q &\propto \int_{constant}^{\infty} dE g(E) E / E^{q/(q-1)} \\ &= \int_{constant}^{\infty} dE g(E) / E^{1/(q-1)} \end{aligned}$$

where  $g(E)$  is the density of states. We can observe that the domain of  $q$  where the two constraints are *finite* (hence mathematically and experimentally well defined) is the *same*; such nice property does *not* occur if we define the energy constraint using  $p_i$  instead of the escort distribution. For example, if we assume the simple case where the high energy approximation is given by  $g(E) \sim E^\delta$  ( $\delta \in \mathbb{R}$ ), the convergence of *both* integrals is guaranteed for  $1/(q-1) - \delta > 1$ , hence  $q < (2+\delta)/(1+\delta)$ . This feature constitutes an ingredient of consistency within the theory, which was not present in the 1988 version. Physically speaking, the internal energy  $U_q$  can be seen as a measure, at a given temperature, of the *width* or *spread* of the distribution  $p_i$  above the lowest possible energy. If we show to a practically-minded scientist say an exponential distribution [ $p(x) \propto e^{-ax}$  with  $a > 0$  for  $x \geq 0$ , and  $p(x) = 0$  for  $x < 0$ ] and a power-law distribution [ $p(x) \propto 1/(1+bx)$  with  $b > 0$  for  $x \geq 0$ , and  $p(x) = 0$  for  $x < 0$ ], and ask him (her) what roughly are the widths, he (she) will promptly check the width at about half value of the maximum for *both* cases, quite independently from the fact that the first moment of the first example is *finite*, whereas it *diverges* for the second one. The expectation value of the energy calculated with the escort distribution  $P_i$ , and *not* with  $p_i$ , precisely is a measure for such width, for *all cases*, independently from what happens for the first moment of the distribution. In other words, this is a *robust* manner for characterizing this particular constraint.

(vii) Let us analyze the same property as in (vi) but for a case where  $\langle x \rangle = 0$ , and consequently we need to refer to the second moment  $\langle x^2 \rangle$ . The already mentioned Lévy-like superdiffusion illustrates such a situation. If we show to the same practically-minded scientist as before say a Gaussian [ $p(x) \propto e^{-ax^2}$ ] and a Lorentzian [ $p(x) \propto 1/(1+bx^2)$ ] distribution, and once again ask him (her) what roughly are the widths, he (she) will once again check the width at about half value of the maximum for *both* cases, quite independently from the fact that the second moment of the first one is *finite*, whereas it *diverges* for the second one. Fixing, for the present Lévy-like example,  $[\int_{-\infty}^{\infty} dx x^2 [p_q(x)]^q] / [\int_{-\infty}^{\infty} dx [p_q(x)]^q]$  precisely characterizes the width for *all* values of  $q$  below 3 (which simultaneously is the upper bound for normalizability!) [25–27]. The Lorentzian example just evoked corresponds to  $q = 2$ .

All the above (i-vii) points were satisfactorily settled in the 1998 version of nonextensive statistical mechanics [14], which is the one that we use since then, hence in the present review. For clarity on the situation, see the Table. However, in spite of the remarkable mathematical consistency of the 1998 version that one might check in this Table, the ultimate geometrical-dynamical understanding of escort distributions is still lacking and is object of current studies (see, for instance, [28]). The situation seems to have some analogy with the following *gedanken* problem. Let us consider a 10 cm linear

object, and let us implement on it the construction of the well known Cantor set, where, at each hierarchy, the central third is deleted. The final object is a fractal with zero one-dimensional Lebesgue measure, and whose Hausdorff dimension is  $d_f = \ln 2 / \ln 3$ . What is the measure of such an object? It is clearly given by  $(10 \text{ cm})^{d_f} = 10^{\ln 2 / \ln 3} \text{ cm}^{\ln 2 / \ln 3} \simeq 4.27 \text{ cm}^{0.63}$ . In other words, we “start” with the entire Lebesgue measure 10 cm, but then we “correct”! The function  $\ln_q W \equiv (W^{1-q} - 1)/(1 - q)$  does something quite analogous: it “starts” with the entire measure  $W$  of the Gibbs  $\Gamma$  space, and then it “corrects”. This would correspond to the fact that, for nonextensive systems, the entire phase space is dynamically *not* nearly entirely occupied (i.e., the system is *not* ergodic), but only a scale-free-like part of it (which depends on the initial conditions) is visited. Consistently, one expects some type of “correction”. It is not impossible that the escort distribution would precisely provide the necessary “correction”. It is perhaps not unworthy to warn the reader that this is but a possible scenario, which remains to be proved.

#### IV. CALCULATING THE INDEX $q$ A PRIORI

For the present theory to be complete, it is obviously necessary to be able to calculate  $q$  *a priori*, i.e., from first principles. Consistently with the thought of Einstein [29], Cohen [30], Baranger [31], and many others, such a basic and crucial task can be accomplished nowhere else than in the analysis of the *microscopic* dynamics (classical, quantum, or others) of the system. However, very helpful connections can also be established at the level of Langevin-like, Fokker-Planck-like, and similar equations, i.e., at the level of *mesoscopic* dynamics. In many occasions, natural and artificial systems whose microscopic or mesoscopic dynamics are either unknown or extremely complex have been addressed. For such systems,  $q$  has been obtained through the fitting of experimental data. We are not addressing these cases here, but more details on several such examples can be found in [15]. Our purpose here is to briefly mention several microscopic and mesoscopic determinations of  $q$  that are available in the literature.

##### A. Determination of $q$ from microscopic dynamics

(i) *Low-dimensional dissipative systems (one- and two-dimensional dissipative maps)*

Let us consider the following one-dimensional dissipative maps ([16,17,32] and references therein):

$$x_{t+1} = 1 - a|x_t|^z \quad (z \geq 1), \quad (47)$$

referred to as the *z-logistic map*, and

$$x_{t+1} = d \cos(\pi|x_t - 1/2|^{z/2}) \quad (z \geq 1). \quad (48)$$

referred to as the  $z$ -periodic map. They belong to the same Feigenbaum-Coulet-Tresser universality classes. More specifically, they share at their edge of chaos (e.g., for  $z = 2$ ,  $a_c = 1.4011\dots$ , and  $d_c = 0.8655\dots$ ), the same Feigenbaum universal constants  $\delta_F(z)$  and  $\alpha_F(z)$ .

Let us also consider ([32] and references therein) a family of one-dimensional dissipative maps which belong to universality classes that are *different* from the just mentioned:

$$\theta_{t+1} = \Omega + [\theta_t - (1/2\pi)\sin(2\pi\theta_t)]^{z/3} \quad (z > 0; 0 < \Omega < 1). \quad (49)$$

referred to as the  $z$ -circle map. The edge of chaos occurs, for  $z = 3$  (the usual case), at  $\Omega_c = 0.6066\dots$ . For all three maps (47), (48) and (49), the following relations have been verified

$$\frac{1}{1 - q_{sen}(z)} = \frac{1}{\alpha_{min}(z)} - \frac{1}{\alpha_{max}(z)} = \frac{(z-1)\ln\alpha_F(z)}{\ln b} \quad (50)$$

where, as mentioned earlier,  $\alpha_{min}$  ( $\alpha_{max}$ ) is the minimal (maximal) value of  $\alpha$  in the  $f(\alpha)$  multifractal function ( $f(\alpha_{min}) = f(\alpha_{max}) = 0$ ), and where  $b = 2$  for maps (47) and (48) and  $b = (\sqrt{5} + 1)/2 = 1.6180\dots$  (*golden mean*) for maps (49).

It is clear that the dissipative Henon map

$$\begin{aligned} x_{t+1} &= 1 - a|x_t|^z + y_t \\ y_{t+1} &= bx_t \quad (0 \leq |b| < 1) \end{aligned} \quad (51)$$

belongs to the same universality class as the logistic map. Checking this case is not really necessary; nevertheless, the numerics have also been directly performed [33]. Consistency with the logistic map has indeed been verified.

To be more explicit,  $q_{sen}(z)$  has been calculated through various independent procedures, namely the sensitivity to the initial conditions [8,9], the multifractal function  $f(\alpha)$  [16], and the finiteness of the entropy production per unit time [10,21]. A fourth method has been used as well, which we describe now. By measuring the shrinking of the Lebesgue measure, Moura et al obtained  $q_{rel}(z)$  for the  $z$ -logistic maps. For example  $q_{rel}(2) \simeq 2.41$ . A connection has been recently obtained [19] between  $q_{sen}$  and  $q_{rel}$ , namely

$$q_{rel}(z, \infty) - q_{rel}(z, W) \propto 1/W^{|q_{sen}(z)|} \quad (52)$$

where  $W$  is the number of cells in which the phase space has been partitioned (the larger  $W$ , the thinner the graining). This equation can be used to simultaneously determine  $q_{sen}(z)$  and  $q_{rel}(z) = q_{rel}(z, \infty)$ .

To illustrate all these methods, let us mention that, since  $\alpha_F(2)$  is known with at least 1018 digits for the logistic map, the corresponding value for

$$q_{sen}(2) = 0.2445\dots$$

is known with the same number of digits. Cases like this one completely disqualify the critical — and, as we see, unfounded — remark that we hear occasionally, namely that  $q$  is no more than a “fitting parameter”.

(ii) *Low-dimensional conservative systems (two- and four-dimensional conservative maps)*

The standard map is defined as follows:

$$\begin{aligned} \theta(t+1) &= p(t) + \frac{a}{2\pi} \sin[2\pi\theta(t)] + \theta(t) \pmod{1}, \\ p(t+1) &= p(t) + \frac{a}{2\pi} \sin[2\pi\theta(t)] \pmod{1} \end{aligned} \quad (53)$$

( $a \in \mathbb{R}$ ,  $t = 0, 1, \dots$ );  $2\pi p$  may in fact be regarded as the angular momentum of a free rotor subject to angle-dependent impulses of strength  $a$  at unit intervals of time.

This two-dimensional map is conservative and symplectic. Also, it is symmetric with regard to  $p = 1/2$ . For  $|a| \gg 1$ , the system is very chaotic (large Lyapunov exponent) and ergodic. When  $a$  approaches  $a_c = 0.9716\dots$  from above, the region around  $p = 1/2$  gradually becomes isolated from the rest of the phase space. Therefore, if we define a kind of “dynamical temperature”  $T(t) \equiv \langle (p(t))^2 \rangle - \langle p(t) \rangle^2$ , a nonzero-measure region exists in phase space where, if we start inside it at  $t = 0$ , we observe two plateaux in  $T(t)$  [34]. The duration of the first plateau diverges when  $a \rightarrow a_c$ . In other words,  $\lim_{t \rightarrow \infty} \lim_{a \rightarrow a_c} T(a, t) \neq \lim_{a \rightarrow a_c} \lim_{t \rightarrow \infty} T(a, t)$ . The existence of such a nonuniform convergence is, as we shall see later on, a distinctive feature which already suggests the possible existence of some degree of nonextensivity. Indeed, anomalies are observed in both the sensitivity to the initial conditions and in the entropy production per unit time. These anomalies appear to be consistent with each other, and they enable the characterization of an index  $q_{sen} < 1$ . However, this index is only an effective one, and depends on  $a$ . For example, for  $a = 0.6$ , we can verify that  $q_{sen} \simeq 0.3$ , but a permanent drift of  $q_{sen}$  is observed towards zero, when we average over many initial conditions all over the entire phase space and consider values of  $a$  gradually approaching zero. The limit  $a = 0$  corresponds of course to the integrable case of the map. Details can be found in [35].

The two-dimensional map has no Arnold diffusion. To verify the influence of dimensionality we also considered two standard maps coupled as follows:

$$\begin{aligned} p_1(t+1) &= p_1(t) + \frac{a_1}{2\pi} \sin[2\pi\theta_1(t)], \\ p_2(t+1) &= p_2(t) + \frac{a_2}{2\pi} \sin[2\pi\theta_2(t)], \\ \theta_1(t+1) &= \theta_1(t) + p_1(t+1) + b p_2(t+1), \\ \theta_2(t+1) &= \theta_2(t) + p_2(t+1) + b p_1(t+1), \end{aligned} \quad (54)$$

where  $a_1, a_2, b \in \mathbb{R}$ ,  $t = 0, 1, \dots$ , and all variables are defined mod 1. If the coupling constant  $b$  vanishes

the two standard maps decouple; if  $b = 2$  the points  $(\theta_1, p_1, \theta_2, p_2) = (0, 1/2, 0, 1/2)$  and  $(1/2, 1/2, 1/2, 1/2)$  are a 2-cycle for all  $(a_1, a_2)$ , hence we preserve, in phase space, the same referential that we had for a single standard map. For a generic value of  $b$ , all relevant present results remain qualitatively the same. Also, we set  $a_1 = a_2 \equiv \tilde{a}$  so that the system is invariant under permutation  $1 \leftrightarrow 2$ . This four-dimensional map is conservative and symplectic, and has Arnold diffusion. It exhibits two plateaus for  $T(t)$ , like the two-dimensional case (with  $\tilde{a}_c = 0$ ). The dependence of  $q_{sen}(\tilde{a})$  is not very different from the two-dimensional case. We are presently studying  $N \gg 1$  coupled maps to verify the effects of the thermodynamic limit. This might or might not stabilize a nontrivial value for  $q_{sen}$ .

Finally, let us mention in this section a connection with quantum chaos, namely for the quantum kicked top Hamiltonian. If we study the time evolution of the fidelity function (scalar product of the wave function of the original Hamiltonian with the wave function of a slightly perturbed Hamiltonian), three types of behaviors are typically observed. (i) In the regions corresponding to regular motion, the fidelity remains roughly constant; (ii) in the regions corresponding to standard chaos (with positive Lyapunov exponents for the associated classical system), the fidelity decreases exponentially with time (before entering the quantum interference region); (iii) at the border of the two regions, a power-law behavior has been recently observed [11] which precisely fits a  $q$ -exponential function. The value of  $q_{rel}$  (associated with the zero perturbation limit) depends on the total momentum value  $J$  of the top. It decreases with increasing  $J$ , and, in the limit  $J \rightarrow \infty$  (classical limit),  $q_{rel} \simeq 2.4$  can be extrapolated [11]. Further analysis is needed and welcome.

*(iii) High-dimensional dissipative systems (many-body dissipative models)*

A specific lattice Lotka-Volterra three-component model has been recently analyzed [36] in terms of entropy production per unit time. For  $d$ -dimensional growth, the following result has been obtained

$$q_{sen} = 1 - 1/d. \quad (55)$$

Details for the  $d = 1, 2$  cases can be found in [36]; the  $d = 3, 4$  cases have been studied by Anteneodo [37]. In fact the microscopic rules of the model are such that the growing droplet has a linear size which grows linearly with time  $t$ . The volume grows then as  $t^d$ . If we apply to this quantity the  $q$ -logarithmic function, we obtain that  $S_q$  scales like  $t$  if  $(1 - q)d = 1$ . From this observation, the result  $q_{sen} = 1 - 1/d$  follows immediately [37]. The possible connection with fractality as  $d$  varies is yet unclear [37].

Boltzmann  $d$ -dimensional Bravais lattice models for the Navier-Stokes equations for incompressible fluids

have been recently revisited [38]. The imposition of the (physically desirable) Galilean invariance of the equations mandates a unique entropy, and that entropy precisely is  $S_q$ , with  $q$  depending on some details of the model ( $q$  is in fact analytically determined by a transcendental equation). For example, for the single-mass single-speed model, it is

$$q = 1 - 2/d. \quad (56)$$

*(iv) High-dimensional conservative systems (many-body Hamiltonian systems)*

Hamiltonian systems is a central topic of statistical mechanics. We shall briefly review here some important results currently available for long-range-interacting classical systems. Let us first focus on the inertial XY ferromagnetic model, characterized by the following Hamiltonian [39,40]:

$$\mathcal{H} = \sum_{i=1}^N \frac{p_i^2}{2} + \sum_{i \neq j} \frac{1 - \cos(\theta_i - \theta_j)}{r_{ij}^\alpha} \quad (\alpha \geq 0), \quad (57)$$

where  $\theta_i$  is the  $i$ -th angle and  $p_i$  the conjugate variable representing the angular momentum (or the rotational velocity since, without loss of generality, unit moment of inertia is assumed). Notice that the summation in the potential is extended to all couples of spins (counted only once) and not restricted to first neighbors; for  $d = 1$ ,  $r_{ij} = 1, 2, 3, \dots$ ; for  $d = 2$ ,  $r_{ij} = 1, \sqrt{2}, 2, \dots$ ; for  $d = 3$ ,  $r_{ij} = 1, \sqrt{2}, \sqrt{3}, 2, \dots$ . The first-neighbor coupling constant has been assumed, without loss of generality, to be equal to unity. This model is an inertial version of the well known XY ferromagnet. Although it does not make any relevant difference, we shall assume periodic boundary conditions, the distance to be considered between a given pair of sites being the smallest one through the  $2d$  possibilities introduced by the periodicity of the lattice. Notice that the two-body potential term has been written in such a way as to have zero energy for the global fundamental state (corresponding to  $p_i = 0, \forall i$ , and all  $\theta_i$  equal among them, and equal to say zero). The  $\alpha \rightarrow \infty$  limit corresponds to only first-neighbor interactions, whereas the  $\alpha = 0$  limit corresponds to infinite-range interactions (a typical Mean Field situation, frequently referred to as the HMF model [39]).

In the limit  $N \rightarrow \infty$ , the quantity  $\tilde{N} \equiv \sum_{i \neq j} r_{ij}^{-\alpha}$  converges to a finite value if  $\alpha/d > 1$ , and diverges like  $N^{1-\alpha/d}$  if  $0 \leq \alpha/d < 1$  (like  $\ln N$  for  $\alpha/d = 1$ ). In other words, the energy is extensive for  $\alpha/d > 1$  and nonextensive otherwise. In the extensive case (here referred to as *short range interactions*; also referred to as *integrable interactions* in the literature), the thermal equilibrium (stationary state attained in the  $t \rightarrow \infty$  limit) is known to be the BG one (see [41]). The situation is much more subtle in the nonextensive case (*long range interactions*). It is this situation that we focus on here. In order to

conform to the most usual writing, we shall from now on replace the Hamiltonian  $\mathcal{H}$  by the following rescaled one:

$$\mathcal{H}' = \sum_{i=1}^N \frac{p_i^2}{2} + \frac{1}{N} \sum_{i \neq j} \frac{1 - \cos(\theta_i - \theta_j)}{r_{ij}^\alpha} \quad (\alpha \geq 0), \quad (58)$$

The molecular dynamical results associated with this Hamiltonian (now artificially transformed into an extensive one for *all* values of  $\alpha/d$ ) can be trivially transformed into those associated with Hamiltonian  $\mathcal{H}$  by re-scaling time (see [40]).

Hamiltonian (58) exhibits in the microcanonical case (isolated system at fixed total energy  $U$ ) a second order phase transition at  $u \equiv U/N = 0.75$ . It has anomalies both above and below this critical point.

Above the critical point it has a Lyapunov spectrum which, in the  $N \rightarrow \infty$  limit, approaches, for  $0 \leq \alpha/d \leq 1$ , zero as  $N^{-\kappa}$ , where  $\kappa(\alpha/d)$  decreases from  $1/3$  to zero when  $\alpha/d$  increases from zero to unity, and remains zero for  $\alpha/d \geq 1$  [40,42]. It has a Maxwellian distribution of velocities [43], and exhibits no aging [44]. Although it has no aging, the typical correlation functions depend on time as a  $q$ -exponential. Diffusion is shown to be of the normal type.

Below the critical point (e.g.,  $u = 0.69$ ), for a nonzero-measure class of initial conditions, a longstanding quasistationary (or metastable) state precedes the arrival to the BG thermal equilibrium state. The duration of this quasistationary state appears to diverge with  $N$  like  $\tilde{N}$  [43,45]. During this anomalous state, there is aging (the correlation functions being well reproduced by  $q$ -exponentials once again), and the velocity distribution is not Maxwellian, but rather approaches a  $q$ -exponential function (with a cutoff at high velocities, as expected for any microcanonical system). Anomalous superdiffusion is shown to exist in this state. The mean kinetic energy ( $\propto T$ , where  $T$  is referred to as the dynamical temperature) slowly approaches the BG value from below, the relaxation function being once again a  $q$ -exponential one. During the anomalous aging state, the zeroth principle of thermodynamics and the basic laws of thermometry have been shown to hold as usual [46,47]. The fact that such basic principles are preserved constitutes a major feature, pointing towards the applicability of thermostatical arguments and methods to this highly nontrivial quasistationary state.

Although none of the above indications constitutes a proof that this long-range system obeys, in one way or another, nonextensive statistical mechanics, the set of so many consistent evidences may be considered as a very strong suggestion that so it is. Anyhow, work is in progress to verify closely this tempting possibility.

Similar observations are in progress for the Heisenberg version of the above Hamiltonian [48], as well as for a  $XY$  model including a local term which breaks the angular isotropy in such a way as to make the model to approach the Ising model [49].

Lennard-Jones small clusters (with  $N$  up to 14) have been numerically studied recently [50]. The distributions of the number of local minima of the potential energy with  $k$  neighboring saddle-points in the configurational phase space can, although not mentioned in the original paper [50], be quite well fitted with  $q$ -exponentials with  $q = 2$ . No explanation is still available for this suggestive fact. Qualitatively speaking, however, the fact that we are talking of very *small* clusters makes that, despite the fact that the Lennard-Jones interaction is not a long-range one thermodynamically speaking (since  $\alpha/d = 6/3 > 1$ ), all the atoms sensibly see each other, therefore fulfilling a nonextensive scenario.

Finally, as a last example of Hamiltonian systems where nonextensive features have been analytically obtained (or numerically observed) we would like to mention anomalous transport in an optical lattice. The distribution of velocities in such system has been recently shown [51] to be, for a specific microscopic model, a  $q$ -distribution with

$$q = 1 + \frac{44E_R}{U_0} \quad (59)$$

where  $E_R$  and  $U_0$  are microscopic parameters of the quantum optical problem.

#### (v) Many-body growth models

A growth model including *preferential* attachment has been recently introduced by Albert and Barabasi [22] as a prototype of emergence of the ubiquitous scale-free networks. At each time step,  $m$  new links are added with probability  $p$ , or  $m$  existing links are rewired with probability  $r$ , or a new node with  $m$  links is added with probability  $1 - p - r$ ; all linkings are done with probability  $\Pi(k_i) = (k_i + 1) / \sum_j (k_j + 1)$ , where  $k_i$  is the number of links of the  $i$ -th node. The exact stationary state distribution of the number  $k$  of links at each site can be written (although apparently not identified by the authors) as  $p(k) \propto e_q^{-k/k_0}$  with

$$q = \frac{2m(2-r) + 1 - p - r}{m(3-2r) + 1 - p - r}, \quad (60)$$

$k_0 > 0$  being an explicit function of  $(p, r, m)$ .

#### (vi) Diffusion in the diluted hypercube

The diluted  $N$ -dimensional hypercube has been considered since long. Its percolation “threshold” is given [52] by  $p_c = \sigma + \frac{3}{2}\sigma^2 + \frac{15}{14}\sigma^3 + \dots$  where  $\sigma \equiv 1/(N-1)$ . Strictly speaking, the quantity  $p_c$  is an effective percolation threshold; indeed the exact percolation threshold clearly is  $p_c = 0$ . Lemke and Almeida [53] have recently studied the model where a random walker diffuses in this structure. More specifically they studied the entropy production for  $N$  increasing up to 23, and a *finite* value was found only for  $q = 1$  at  $p = 1$  (full hypercube), and for  $q = 0.056$  at  $p = p_c$ .

## B. Determination of $q$ from mesoscopic dynamics

### (i) Nonlinear Fokker-Planck equations (Correlated anomalous diffusion)

In a variety of physical situations [26], it is appropriate to consider the following nonlinear Fokker-Planck-like equation (sometimes referred to as the *Porous Medium Equation*):

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}[F(x)p(x,t)] + D\frac{\partial^2[p(x,t)]^\nu}{\partial x^2} \quad (\nu \in \mathcal{R}). \quad (61)$$

If we assume that at  $t = 0$  we have the paradigmatic (and quite usual) distribution  $p(x,0) = \delta(x)$ , it can be shown [26] that, for  $F(x) = k_1 - k_2x$  ( $k_1 \in \mathcal{R}$ ;  $k_2 \geq 0$ ) and all  $(x,t)$ , the (stable) solution is given by

$$p(x,t) \propto e_q^{-\beta(t)[x-x_M(t)]^2} \quad (q < 3), \quad (62)$$

where  $\beta(t)$  and  $x_M(t)$  are smooth explicit functions of  $t$ , and

$$q = 2 - \nu \quad (q < 3). \quad (63)$$

### (ii) Fractional-derivative Fokker-Planck equations (Lévy anomalous diffusion)

It is known [54] that  $p(x) \propto e_q^{-\beta x^2}$  ( $q < 3$ ) optimizes

$$S_q = \frac{1 - \int dx [p(x)]^q}{q - 1} \quad (64)$$

under appropriate constraints. If we convolute  $N$  times  $p(x)$  ( $N \rightarrow \infty$ ), we approach a Gaussian distribution if  $q < 5/3$  and a Lévy  $L_{\gamma_L}(x)$  one if  $5/3 < q < 3$ . Lévy distributions are the solutions of the following equation  $\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^{\gamma_L} p(x,t)}{\partial x^{\gamma_L}}$  ( $0 < \gamma_L < 2$ ). The index  $\gamma_L$  of these Lévy distributions is related to  $q$  as follows:

$$q = \frac{\gamma_L + 3}{\gamma_L + 1} \quad (5/3 < q < 3). \quad (65)$$

In general, the Lévy distribution and the  $q$ -Gaussian coincide only asymptotically (the tails). The unique exception is for  $q = 2\gamma_L = 2$ , in which case they are identical.

### (iii) Nonlinear fractional-derivative Fokker-Planck equations

The equation

$$\frac{\partial p(x,t)}{\partial t} = D\frac{\partial^\gamma [p(x,t)]^\nu}{\partial x^\gamma} \quad (\nu \in \mathcal{R}; \gamma \in \mathcal{R}). \quad (66)$$

is an interesting one. The  $\gamma = 2$  particular instance reproduces the situation addressed with Eq. (61); the

$\nu = 1$  particular instance reproduces the partial differential equation above Eq. (65). We do not know the solution for the generic case. However, we do know it for the (new) case  $\nu = (2 - \gamma)/(1 + \gamma)$  [55]. It is given by

$$q = \frac{\gamma + 3}{\gamma + 1} = \frac{5 + 2\nu}{3} \quad (67)$$

with  $0 < \gamma, \nu < 2$ . As for the Lévy case, it is only asymptotically that the solution coincides with the  $q$ -Gaussian function. More, and more complex, situations along these lines can be found in [56].

### (iv) Anomalous Langevin equations

We may consider the following Langevin-like equation [27]

$$\dot{x} = -\gamma x|x|^{2(s-1)} + x|x|^{s-1}\xi(t) + \eta(t) \quad (s > 0), \quad (68)$$

where  $\xi(t)$  and  $\eta(t)$  are independent and Gaussian-distributed zero-mean white noises, satisfying

$$\langle \xi(t)\xi(t') \rangle = 2M\delta(t - t') \quad (M \geq 0) \quad (69)$$

and

$$\langle \eta(t)\eta(t') \rangle = 2A\delta(t - t') \quad (A > 0). \quad (70)$$

$M$  and  $A$  stand for *multiplicative* and *additive* respectively. If  $\gamma \geq M(1 - s)$ , the distribution corresponding to the stationary state is given by

$$P(x) \propto e_q^{-\frac{(\gamma/s)+M}{2A}|x|^{2s}}, \quad (71)$$

with

$$q = \frac{(\gamma/s) + 3M}{(\gamma/s) + M}. \quad (72)$$

Notice that, interestingly enough, the case  $s = M = 1$  reproduces the results indicated in Eqs. (65) and (67).

Another interesting Langevin-like case is that of dichotomic colored noise addressed in [57]. One may consider the following stochastic differential equation:

$$\frac{dV}{dt} = -\gamma V + \xi(t), \quad (73)$$

where  $\xi(t)$  is a dichotomous noise of values  $\pm a$ , satisfying

$$\langle \xi(t + \tau)\xi \rangle = a^2 e^{-2\lambda|\tau|}. \quad (74)$$

The stationary state probability is, although not written in this way in the original paper [57], given by  $p(V) \propto e_q^{-\beta V^2}$  with

$$q = \frac{1 - 2\gamma/\lambda}{1 - \gamma/\lambda} \quad (\gamma/\lambda \leq 1), \quad (75)$$

$\beta$  being given by some smooth function of the model parameters.

(v) *Phenomenological approach of pressure fluctuations in multiphase flow*

A granular system has been experimentally studied [58] which consists of bubbling fluidized beds set in motion by a vertical gas stream. In addition to the experiments themselves, a phenomenological approach has been advanced. It provides, for the distribution of the pressure fluctuations, a  $q$ -Gaussian with

$$q = 1 + \frac{1}{\tau - 1/2} \quad (\tau > 3/2), \quad (76)$$

where  $\tau$  is a phenomenological parameter of the bubble size distribution. We must warn the reader that the authors argue [58] that, from their standpoint, no deep connection exists with nonextensive statistical mechanics. We do not necessarily agree with their interpretation. But it is not the aim of this short note to address this particular point in depth. We just refer the existence of this mathematical result in connection with those experiments.

(vi) *Scalar granular gases*

Baldassarri et al [59] have numerically and analytically discussed some isolated one-dimensional granular gases in a recent work. A restitution coefficient  $r \in [0, 1)$  has been assumed for two-body collisions. The velocity distribution in the asymptotic cooling regime of a specific pseudo-Maxwell (Ulam's) model, has been shown to be a distribution, which, although the authors apparently have not identified it as such, precisely is a  $q$ -Gaussian with  $q = 3/2$ .

## V. TOWARDS A GENERALIZED NONEXTENSIVE STATISTICAL MECHANICS

A statistical mechanics can be seen as basically consisting in the appropriate choice of a physical entropic functional  $S(\{p_i\})$ . Once this is done, we must consistently choose the constraints ( $\sum_{i=1}^W p_i = 1$ , and possibly others) to be used for describing various physical situations (isolated, or in equilibrium with an infinitely large reservoir, or others). We must then extremize (typically maximize) the entropy under these constraints, in order to obtain the probability distribution corresponding to the stationary state, whenever it exists. Relaxation phenomena towards the distribution associated with the stationary state(s) can be characterized through  $S(\{p_i(t)\})$ , among others. The stationary state entropy is of course given by  $S(\{p_i(\infty)\})$ . The experimental information can provide (direct or indirect) evidence on  $\{p_i(t)\}$ , and on  $S(t) = S(\{p_i(t)\})$ . But establishing the functional  $S(\{p_i\})$

itself demands the knowledge of  $S(t)$  for almost all possible trajectories. So, in practice, the functional  $S(\{p_i\})$  is not deduced or measured, but found heuristically. This functional is in fact a mathematical construct which, for a variety of macroscopic physical quantities, can essentially replace (as the genius of Boltzmann showed to us) the detailed knowledge implied at the microscopic dynamical level.

The microscopic dynamics of an isolated system might be such that, for practically all possible initial conditions and as time increases, it quickly evolves towards an *uniform* occupation of phase space (either of the entire phase space, as for the  $N \rightarrow \infty$   $d = 3$  short-range Heisenberg ferromagnet above its critical total internal energy, or of a simple and well defined part of it, as for the same ferromagnet *below* its critical internal energy, once a specific symmetry has been broken). The system is then said *ergodic*, and Boltzmann's basic equal-probability hypothesis indeed applies. The entropy to be used in such a case is widely known to be  $S_{BG}$ .

What happens, however, if the system has more complex dynamics, such that it tends, when isolated, to occupy phase space in a nontrivial nonuniform manner? This is where generalizations of BG statistical mechanics possibly become a natural way out. This imposes, in our opinion, the need of physical entropies *different* from the BG one.

The nonextensive entropy  $S_q(\{p_i\})$

(i) satisfies ( $\forall q > 0$ ), among others, three mathematical properties (namely concavity, experimental robustness, and finiteness of the entropy production per unit time) that are *by no means trivial* to satisfy (e.g., Renyi entropy satisfies none of these  $\forall q > 0$ ),

(ii) yields a stationary distribution (the  $q$ -exponential function, which is asymptotically a power-law) which is ubiquitously found in natural and artificial systems, and

(iii) is in principle completely determined once its entropic index  $q$  is calculated from microscopic or mesoscopic dynamics.

It is therefore allowed to consider  $S_q$  as a quite strong candidate for physically extending the applicability of the methods of BG statistical mechanics to a specific class of dynamical systems among those who do *not* comply with the usual ergodicity requirements. The rigorous characterization of this class is yet to be better understood, but we already dispose of a good basis for conjecturing that it concerns systems that mix slowly in phase space, yielding a long-standing (multi) fractal-like occupation of it (possibly with a *scale-free* structure, as it is nowadays called in the field of networks). Such is the case of various dissipative low-dimensional nonlinear dynamical systems at their edge of chaos, conservative low-dimensional nonlinear dynamical systems close to the frontier between chaos and integrability, (possibly) long-range-interacting Hamiltonian systems, and others.

Are we compelled to stop here, or can we hope to similarly cover even more complex dynamics, eventually

yielding stationary states (or sequences of quasistationary states) whose phase space structure is even more complex than the one just mentioned? The bottom line is that we see no basic reason for stopping. However, a highly nontrivial problem must be solved before further progressing along this line, namely the proposal of an appropriate entropy functional. There are at least two lines of thought (hereafter referred to as the *differential equation path*, and the *superstatistics path*) that converge onto a solution of such a task. Let us next briefly mention both.

### A. The differential-equation path

We have seen that BG and nonextensive statistical mechanics are deeply related to the differential equations (5) and (26) respectively. Can we unify them in such a way as to have a crossover between one and the other, depending on the value of the independent variable  $x$ ? Yes, we can, and this has been advanced in 1999 [60] for re-association in folded proteins, and recently used for a phenomenological theory for the flux of cosmic rays [61]. The associated differential equation we propose is as follows:

$$\frac{dy}{dx} = a_1 y + (a_q - a_1) y^q \quad (y(0) = 1) , \quad (77)$$

hence

$$y = \left[ 1 - \frac{a_q}{a_1} \left( 1 - e^{(1-q)a_1 x} \right) \right]^{\frac{1}{1-q}} \quad (78)$$

Eq. (77) recovers Eq. (26) if we consider  $a_1 = 0$ . It recovers Eq. (5) if we consider  $a_q = a_1$ , and also if we just consider  $q = 1$ . The existence of this last (convenient) possibility is the reason for which we presented the generalization in the form of Eq. (77), rather than in the totally equivalent form  $\frac{dy}{dx} = a_1 y + a'_q y^q$ . An interesting situation occurs for the case  $q < 1$  with  $0 < a_1 \ll a_q$ , as well as for the case  $q > 1$  with  $a_q \ll a_1 < 0$ . For  $0 \leq x \ll x_{\text{crossover}} \equiv [1/[(1-q)a_q]]$  we have  $y \sim e_q^{a_q x}$  (nonextensive statistical mechanics), whereas for  $x \gg x_{\text{crossover}}$  we have  $y \sim e^{a_1 x}$  (BG statistical mechanics).

We can further generalize Eq. (77) in a natural way, namely through

$$\frac{dy}{dx} = a_{q'} y^{q'} + (a_q - a_{q'}) y^q \quad (y(0) = 1) , \quad (79)$$

whose solution is a combination of hypergeometric functions (see [60]). As before, an interesting situation occurs for the case  $q < q' < 1$  with  $0 < a_{q'} \ll a_q$ , as well as for the case  $q > q'$  with  $a_q \ll a_{q'} < 0$ . For  $0 \leq x \ll x_{\text{crossover}} \equiv [(1-q)a_q]^{\frac{1-q'}{q'-q}} / [(1-q')a_{q'}]^{\frac{1-q}{q'-q}}$  we have  $y \sim e_q^{a_q x}$  ( $q$ -nonextensive statistical mechanics), whereas for  $x \gg x_{\text{crossover}}$  we have  $y \sim e_{q'}^{a_{q'} x}$  ( $q'$ -nonextensive statistical mechanics).

Before closing this subsection, let us make a remark concerning the history of physics. It seems quite reasonable to discuss the time dependence of quantities like the sensitivity to initial conditions and the relaxation of physical quantities in terms of a differential equation. But to do so for the energy distribution associated with a stationary state seems, as earlier mentioned, rather amazing. To make such procedure somewhat more acceptable it might be interesting at this point to review a few equations of Planck's October 1900 paper [2]. He writes the following two equations:

$$\frac{d^2 S}{dU^2} = \frac{\alpha}{U(\beta + U)} , \quad (80)$$

and

$$\frac{dS}{dU} = \frac{1}{T} , \quad (81)$$

where  $S$  is the entropy,  $U$  the internal energy,  $T$  the absolute temperature,  $\alpha$  and  $\beta$  two coefficients to be fixed. The first of these two equations was heuristic. If we replace the second equation into the first one, we obtain

$$\frac{d}{dU} \left( \frac{1}{T} \right) = \frac{\alpha}{U(\beta + U)} \quad (82)$$

hence

$$\frac{dU}{d(1/T)} = \frac{\beta}{\alpha} U + \frac{1}{\alpha} U^2 . \quad (83)$$

It happens that this equation precisely is the  $(q', q) = (1, 2)$  particular case of our Eq. (77). Planck, in fact, does not write our present Eqs. (82) and (83). After writing our present Eqs. (80) and (81), he jumps (taking into account the density of states) to the writing of the celebrated law for the radiation of the black body. Although not always noticed, the bosonic nature of photons (i.e., the special constraint implied by the symmetrization of the quantum wavefunctions) makes that Planck's law does undergo through a crossover from a power law ( $q = 2$ ) at low energies to an exponential law ( $q = 1$ ) at high energies.

### B. The superstatistics path

It was noticed in 2000 by Wilk and Włodarczyk [62] that the  $q$ -exponential distribution can be written as a Laplace transform involving the BG exponential weight. More precisely,

$$e_q^{-\beta_q E_i} = \int_0^\infty d\beta e^{-\beta E_i} f_q(\beta) , \quad (84)$$

where  $f_q(\beta)$  is the  $\chi^2$  (also called the Gamma) distribution. They further noticed that

$$q = \frac{\langle \beta^2 \rangle}{\langle \beta \rangle^2} . \quad (85)$$

where  $\langle \cdots \rangle \equiv \int_0^\infty d\beta (\cdots) f_q(\beta)$ . This observation was further developed by Beck [63]. The idea that was emerging was the possible interpretation of nonextensive statistics as a kind of average of BG statistics, where the temperature (or scaled coupling constants) itself would be a stochastic variable. Such mathematical relations were in fact already present in the so called Hilhorst (see [64]), Prato [65] and Lenzi [66] formulae, however without any particular physical interpretation.

The next step was recently done by Beck and Cohen when they proposed, through a Laplace transform, their superstatistics [67]. They generalized Eq. (84) into

$$B(E_i) = \int_0^\infty d\beta e^{-\beta E_i} f(\beta), \quad (86)$$

where  $f(\beta)$  is a quite general distribution ( $\int_0^\infty d\beta f(\beta) = 1$ ). They consistently generalized Eq. (85) into

$$q_{BC} \equiv \frac{\langle \beta^2 \rangle}{\langle \beta \rangle^2}, \quad (87)$$

where  $\langle \cdots \rangle \equiv \int_0^\infty d\beta (\cdots) f(\beta)$  (the subindex  $BC$  stands for Beck-Cohen and was introduced in [68] to avoid confusion with  $q$ ). Clearly, if  $f(\beta) = f_q(\beta)$  we have that  $q_{BC} = q$ .

Of course, having a weight like that defined in Eq. (86) is necessary but not sufficient for having a statistical mechanics. We also need an entropy functional, and a consistent manner of writing the constraints such as the energy and similar ones. The optimization of this entropy is expected to be extremal for the superstatistical weight  $B(E_i)$ . Such an entropy, from now on noted  $S_G(\{p_i\})$  ( $G$  stands for *generalized*), was in fact proposed in [68]. It has an univocal relation with  $f(\beta)$ , i.e., one and only one functional  $S_G$  corresponds to each admissible  $f(\beta)$ . Of course,  $f(\beta) = f_q(\beta)$  implies and is implied by  $S_G = S_q$ . This sensibly general entropy  $S_G$  is concave by construction. It also happens to be experimentally robust, as proved in [69]. Such a convergence of two non-trivial properties (concavity and robustness) is certainly very satisfactory, and constitutes an argument favorable to considering  $S_G$  as a physical entropy (i.e., useful for thermal physics), and not just as one more theoretical information measure, among dozens that exist in the literature of cybernetics, control theory, and other applied sciences.

### C. Connecting the differential-equation and the superstatistics paths

The solution of Eq. (79) clearly corresponds to a particular  $B(E_i)$ , from now on noted  $B_{qq'}(E_i)$ , and to its associated  $f(\beta)$ , from now on noted  $f_{qq'}(\beta)$ . In particular, we have that  $B_{q1}(E_i)$  equals the function given in Eq. (78) with  $(a_1, a_q, x) \rightarrow (-\beta_1, -\beta_q, E_i)$ . More particularly,

$$B_{21}(E) = \frac{1}{1 - \frac{\beta_2}{\beta_1}(1 - e^{\beta_1 E})}. \quad (88)$$

If we normalize this function in order to get the probability distribution, take the limit  $\beta_2/\beta_1 \rightarrow \infty$ , and multiply by the  $d = 3$  photonic density of states, we recover, as already discussed, Planck's law.

More details about  $B_{qq'}(E_i)$ ,  $f_{qq'}(\beta)$ , and their application to the recent high-precision experimental (by Bodenschatz et al [70]) and computational (by Gotoh et al [71]) results for fully developed turbulence, can be found in [72]. Finally, the logical structure of the successive generalizations of *classical* BG statistical mechanics presented in this paper is indicated in Fig. 2. The case of the Planck black-body distribution *outside* BG statistical mechanics in the figure deserves a clarification. It is well known that, photons being bosons, this distribution naturally belongs to the *quantum* version of BG statistical mechanics, more precisely to Bose-Einstein statistics. To be more explicit, *classical* BG statistics is contained inside *quantum* BG statistics. Bose-Einstein statistics (as well as Fermi-Dirac statistics) is also contained *inside* quantum BG statistics though *outside* from classical BG statistics. All these distributions (classical BG, Bose-Einstein and Fermi-Dirac) can obviously be conceived as particular cases of the Beck-Cohen superstatistics since only a Laplace transform is involved (see Eq. (86)). What is special about Planck law is that it can be *also* conceived as a simple particular case of the special instance of superstatistics which simultaneously is a solution of Eq. (77). In other words, the celebrated distribution can also be seen as a crossover between classical BG statistics (i.e., classical  $q = 1$  statistics) and classical  $q = 2$  statistics. Summarizing, it is in this *specific sense* that this distribution is, as indicated in Fig. 2, *outside* classical BG statistics, and *inside* superstatistics.

## VI. CRITIQUES

As the history of sciences plethorically shows to us, every possible substantial progress in the foundations of any science is accompanied by controversies. This is a common and convenient mechanism for new ideas to be checked and better understood by the scientific community. There is absolutely no reason to expect that statistical mechanics would be out of it. Quite on the contrary [73], given the undeniable fact that entropy is one among the most subtle and rich concepts in physics. Indeed, as eloquently commented by Nicolis and Daems [73], "*It is the strange privilege of statistical mechanics to stimulate and nourish passionate discussions related to its foundations [...]*". We believe that some space dedicated here to such issues might well be useful at this stage. Therefore, let us address one by one some recent critiques that we are aware of. In fact they have all been replied in specific papers which we indicate here case by case for the interested reader. In what follows, for

economy of space, we shall restrict to a brief review of what we understand to be the main focus of each criticism, and what we believe to be the main reason for its rebuttal.

(a) *Vollmayr-Lee and Luijten critique:* [74]

Vollmayr-Lee and Luijten (VLL) presented in 2001 [74] a critique to nonextensive statistical mechanics. They consider a Kac-potential approach of nonintegrable interactions. They consider a  $d$ -dimensional classical fluid with two-body interactions exhibiting a hard core as well as an attractive potential proportional to  $r^{-\alpha}$  with  $0 \leq \alpha/d < 1$  (logarithmic dependence for  $\alpha/d = 1$ ; VLL use the notation  $\tau \equiv \alpha$ ). In their approach, they also include a Kac-like long-distance cutoff  $R$  such that no interactions exist for  $r > R$ , and then discuss the  $R \rightarrow \infty$  limit. They show that the exact solution within Boltzmann-Gibbs statistical mechanics is possible and that – no surprise (see VLL Ref. [12] and references therein) – it exhibits a mean field criticality. Moreover, the authors argue that very similar considerations hold for lattice gases,  $O(n)$  and Potts models.

VLL state “*Our findings imply that, contrary to some claims, Boltzmann-Gibbs statistics is sufficient for a standard description of this class of nonintegrable interactions.*”, and also that “*we show that nonintegrable interactions do not require the application of generalized  $q$ -statistics.*”. In our opinion, these statements may misguide the reader. The critique was rebutted in [75], whose main points are summarized here. Indeed, the VLL discussion, along traditional lines, of their specific Kac-like model only exhibits that Boltzmann-Gibbs statistical mechanics is – as more than one century of brilliant successes guarantees! – *necessary* for calculating, *without doing time averages*, a variety of *thermal equilibrium* properties; by no means it proves that it is *sufficient*, as we shall soon clarify. Neither it proves that wider approaches such as, for instance, nonextensive statistical mechanics (VLL Refs. [6,31] and present [12–14]), or any other similar formalism that might emerge, are not required or convenient. The crucial point concerns *time*, a word that nowhere appears in the VLL paper. The key role of  $t$  has been emphasized in several occasions, for instance in Fig. 4 of VLL Ref. [31]. For integrable or short-range interactions (i.e., for  $\alpha/d > 1$ ), we expect that the  $t \rightarrow \infty$  and  $N \rightarrow \infty$  limits are commutable in what concerns the equilibrium distribution  $p(E)$ ,  $E$  being the total energy level associated with the macroscopic system. More precisely, we expect naturally that (excepting for the density of states)

$$p(E) \equiv \lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} p(E; N; t) = \lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} p(E; N; t) \propto \exp[-E/kT] \quad (\tau/d > 1) \quad (89)$$

if the system is in thermal equilibrium with a thermostat at temperature  $T$ . In contrast, the system is expected to behave in a more complex manner for nonintegrable

(or long-range) interactions, i.e., for  $0 \leq \alpha/d \leq 1$ . In this case, no generic reason seem to exist for the  $t \rightarrow \infty$  and  $N \rightarrow \infty$  limits to be commutable, and consistently we expect not necessarily equal results. The simplest of these results (which is in fact the one to be associated with the VLL paper, although therein these two relevant limits and their ordering are not mentioned) is, as we shall soon further comment,

$$\lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} p(E; N; t) \propto \exp[-(E/\tilde{N})/(kT/\tilde{N})] . \quad (90)$$

$\tilde{N} \equiv [N^{1-\alpha/d} - \alpha/d]/[1 - \alpha/d]$  has been introduced in order to stress the facts that *generically*

(i)  $E$  is *not* extensive, i.e., is *not* proportional to  $N$  but is instead  $E \propto N\tilde{N}$  [more precisely,  $E$  is extensive if  $\alpha/d > 1$  (see [76] and VLL Refs. [4,5]), and it is nonextensive if  $0 \leq \alpha/d \leq 1$ ]; and

(ii)  $T$  needs, in such calculation, to be rescaled (a feature which is frequently absorbed in the literature by artificially size-rescaling the coupling constants of the Hamiltonian), in order to guarantee nontrivial *finite* equations of states. Of course, for  $\alpha = 0$ , we have  $\tilde{N} = N$ , which recovers the traditional Mean Field scaling.

But, *depending on the initial conditions*, which determine the time evolution of the system if it is assumed isolated, quite *different* results can be obtained for the ordering  $\lim_{t \rightarrow \infty} \lim_{N \rightarrow \infty} p(E; N; t)$ . This fact has been profusely detected and stressed in the related literature (see, for instance, VLL Ref. [31], present Refs. [39,43–45,77–80] and references therein). Unfortunately, this important fact has been missed in the VLL critique. Such metastable states can *by no means* be described within BG statistical mechanics. Even more, as shown in Section IV, there is nowadays increasing evidence that they might be intimately related to nonextensive statistical mechanics. In any case, it is plain that, for such long-range Hamiltonians, BG statistics is necessary but not sufficient, in contrast with the VLL statements.

(b) *Nauenberg critique:* [81]

Some line of critique concerns whether the zeroth principle of thermodynamics and thermometry are consistent with nonextensive statistical mechanics. Such questioning is by no means new: a couple of dozens of papers are available in the literature which address this important point. It has been recently raised once again, this time by Nauenberg [81]. He concludes, among many other critiques, that it is not possible to have thermalization between systems with different values of  $q$ . It appears to be exactly the opposite which is *factually* shown in [46]. His critique is rebutted in [46]. One of the crucial points that is unfortunately missed in [81], concerns discussion of “weak coupling” in Hamiltonian systems. Indeed, if we call  $c$  the coupling constant associated with long range interactions (i.e.,  $0 \leq \alpha/d \leq 1$ ), we have that  $\lim_{N \rightarrow \infty} \lim_{c \rightarrow 0} c\tilde{N} = 0$ , whereas  $\lim_{c \rightarrow 0} \lim_{N \rightarrow \infty} c\tilde{N}$

diverges. No such anomaly exists for short-range interactions (i.e.,  $\alpha/d > 1$ ). Indeed, in this simpler case, we have that  $\lim_{N \rightarrow \infty} \lim_{c \rightarrow 0} c\tilde{N} = \lim_{c \rightarrow 0} \lim_{N \rightarrow \infty} c\tilde{N} = 0$ . The nonuniform convergence that, for long-range interactions, exists at this level possibly is related to the concomitant nonuniform convergence associated with the  $t \rightarrow \infty$  and  $N \rightarrow \infty$  limits discussed previously in this paper. All these subtleties are not mentioned in [81].

(c) *Luzzi, Vasconcellos and Ramos critique*: [82]

Another line of critique concerns the “physicality” of  $S_q$  (see [82]). Or whether it could exist a “physical” entropy different from  $S_{BG}$ . Since such issues appear to be of a rather discursive/philosophical nature, we prefer to put these critiques on slightly different, more objective, grounds. We prefer to ask, for instance, (i) whether  $S_q$  is useful in theoretical physics in a sense similar to that in which  $S_{BG}$  undoubtedly is useful; (ii) whether  $q$  necessarily is a fitting parameter, or whether it can be determined a priori, as it should if we wish the present theory to be a complete one; (iii) whether there is no other way of addressing the thermal physics of the anomalous systems addressed here, very specifically whether one could not do so by just using  $S_{BG}$ ; (iv) whether  $S_q$  is special in some physical sense, or whether it is to be put on the same grounds as the twenty or thirty entropic functionals popular in cybernetics, control theory, and information theory generally speaking. Such questions have received answers in [83–88] and elsewhere. (i) The usefulness of this theory seems to be answered by the large amount of applications it has already received, and by the ubiquity of the  $q$ -exponential form in nature. (ii) The a priori calculation of  $q$  from microscopic dynamics has been specifically illustrated in Section III. (iii) The optimization of  $S_q$ , as well as of almost any other entropic form, with a few constraints has been shown in [84] to be equivalent to the optimization of  $S_{BG}$  with an infinite number of appropriately chosen constraints. Therefore we could restrain to the use of  $S_{BG}$  if we absolutely wanted that, similarly to the fact that, instead of using the extremely convenient Keplerian ellipse for the planetary orbits, we could equivalently use an infinite number of Ptolemaic epicycles. It is however appreciably much simpler to represent a complex structure of constraints into the single index  $q \neq 1$  (in analogy with the fact that the ellipticity of a Keplerian orbit can be simply specified by a single parameter, namely the eccentricity of the ellipse). (iv) The entropy  $S_q$  shares with  $S_{BG}$  an impressive set of important properties, which namely includes concavity, stability and finiteness of the entropy production per unit time,  $\forall q > 0$ . The difficulty of simultaneously satisfying all these three properties can be measured by the fact that the Renyi entropy (usefully used in the geometric characterization of multifractals) satisfies none of them for all  $q > 0$ . Such features point  $S_q$  as being very special for thermostistical purposes.

(d) *Zanette and Montemurro critique I*: [89]

In a recent paper [89], Zanette and Montemurro re-analyze the molecular dynamics approach and results presented in [43] for the infinitely-long-range interacting planar rotators already discussed here. They especially focus on the time dependence of the temperature  $T(t)$  defined as the mean kinetic energy per particle. For total energy slightly below the second-order critical point and a non-zero-measure class of initial conditions, a long-standing nonequilibrium state emerges before the system achieves the terminal BG thermal equilibrium. When  $T(t)$  is plotted, as done by virtually all authors, by using a linear scale for  $T$  and a logarithmic scale for  $t$ , an inflection point exists. If we call  $t_{crossover}$  the value of  $t$  at which the inflection point is located, it has been repeatedly verified numerically by various authors, including Zanette and Montemurro [89], that  $\lim_{N \rightarrow \infty} t_{crossover}(N)$  diverges. Therefore, if the system is very large (in the limit  $N \rightarrow \infty$ , mathematically speaking) it remains virtually for ever in the anomalous state, currently called by many authors, *quasi-stationary state* or *metastable state*. Zanette and Montemurro point out (correctly) that, if a linear scale is used for  $t$ , the inflection point disappears. From this, these authors conclude that this well known metastable state is but a kind of mathematical artifact, and no physically relevant quasi-stationarity exists. It is like if the high-to-low energies crossing occurring, at a given temperature, in Fermi-Dirac statistics had no physical meaning! Indeed, if instead of using the linear scale for the energies we were to use a faster scale (e.g., an exponential scale), the well known inflection point will surely disappear. Nevertheless, there is no point to conclude from this that the textbook crossing in Fermi-Dirac statistics is but a mathematical artifact. In fact, *any* inflection point on *any* curve will disappear by sufficiently “accelerating” the abscissa. Coming back to the system of rotators, what indeed appears to happen is that, for increasingly large  $N$ ,  $T(t)$  remains constant, and *different from the BG value*, within an error bar which appears to *vanish* in the  $N \rightarrow \infty$  limit. This effect appears in an even more pronounced way because of a slight minimum that  $T(t)$  presents just before going up to the BG value. This intriguing minimum had already been observed in [43] and has been detected with higher precision in [89]. A partial rebuttal of this critique is presented in [77].

(e) *Zanette and Montemurro critique II*: [90]

Soon after the previous one, Zanette and Montemurro advanced a second critique [90] objecting the validity of nonextensive statistical mechanics for thermodynamical systems. This line of critique concerns the ubiquity of the  $q$ -exponential form as a stable law in nature. The argument essentially goes that only Gaussians and Lévy distributions would be admissible, because of the

respective central limit theorems. Such question has been addressed long ago in [55] (ignored in [90]), and once again more recently in [91], as a rebuttal to [90]. The answer basically reminds that the time stability of the distribution only has a mathematically well defined meaning if we *also* specify the time evolution or time composition law. The just mentioned central limit theorems *only* apply for convolution time evolution, certainly not for nonmarkovian evolutions as those illustrated in Section IV.

(f) *Zanette and Montemurro critique III*: [92]

Soon after the second critique, Zanette and Montemurro advanced a third one [92]. This time the objection addresses the rest of the systems, i.e., the non thermodynamical ones. It is argued by these authors that non thermodynamical applications of nonextensive statistics are ill-defined, essentially because of the fact that *any* probability distribution can be obtained from the nonextensive entropy  $S_q$  by conveniently adjusting the constraint used in the optimization. We argue here that, since it is well known to be so for *any* entropic form and, in particular, for the Boltzmann-Gibbs entropy  $S_{BG}$ , the critique brings absolutely no novelty to the area. In other words, it has *nothing* special to do with the nonextensive entropy  $S_q$ . In defense of the usual simple constraints, typically averages of the random variable  $x_i$  or of  $x_i^2$  (where  $x_i$  is to be identified according to the nature of the system), we argue, *and this for all entropic forms*, that they can hardly be considered as arbitrary, as Zanette and Montemurro seem to consider. Indeed, once the natural variables of the system have been identified, the variable itself and its square obviously are the most basic quantities to be constrained. Such constraints are used in hundreds (perhaps thousands) of useful applications outside (and also inside) thermodynamical systems, along the information theory lines of Jaynes and Shannon, and more recently of A. Plastino and others. And this is so for  $S_{BG}$ ,  $S_q$ , and any other entropic form. If, however, other quantities are constrained (e.g., an average of  $x^\sigma$  or of  $|x|^\sigma$ ) for specific applications, it is clear that, at the present state-of-the-art of information theory, *and for all entropic forms*, this must be discussed case by case. The full rebuttal of this critique can be found in [93].

(g) *About ordinary differential equations*:

The remarks in Sections I.A and II.A related to ordinary differential equations might surprise some readers, hence deserve some clarification. Indeed, in virtually all the textbooks of statistical mechanics, functions such as the energy distribution at thermal equilibrium are discussed *using a variational principle*, namely referring to the entropy functional, and *not using ordinary differential equations* and their solutions. In our opinion, it is so *not* because of some basic (and unknown) principle of

exclusivity, but rather because the first-principle dynamical origin of the BG factor still remains, mathematically speaking, at the status of a *dogma* [94]. Indeed, to the best of our knowledge, no theorem yet exists which establishes the necessary and sufficient conditions for being valid the use of the celebrated BG factor. Nevertheless, one must not forget that it was precisely through a differential equation that Planck heuristically found, as described in his famous October 1900 paper [2], the black-body radiation law. It was only in his equally famous December 1900 paper that he made the junction with the Boltzmann factor by assuming the — at the time, totally bizarre — hypothesis of discretized energies.

A further point which deserves clarification is *why* have we *also* interpreted the linear ordinary differential equation in Section I.A as providing the typical time evolution of both the sensitivity to the initial conditions and the relaxation of relevant quantities. Although the bridging was initiated by Krylov [1], the situation still is far from completely clear on mathematical grounds. However, intuitively speaking, it seems quite natural to think that the sensitivity to the initial conditions is precisely what makes the system to relax to equilibrium, and therefore opens the door for the BG factor to be valid. In any case, although some of the statements in Section I.A are (yet) not proved, this by no means implies that they are generically false. Furthermore, they provide what we believe to be a powerful metaphor for generalizing the whole scheme into the nonlinear ordinary differential equations discussed in Section II.A. Interestingly enough, the  $q$ -exponential functions thus obtained have indeed proved to be the correct answers for a sensible variety of specific situations reviewed in the rest of the present paper, and this for *all three* interpretations as energy distribution for the stationary state, time evolution of the sensitivity to the initial conditions, and time evolution of basic relaxation functions.

## VII. FINAL COMMENTS

We all know that the concepts of energy and entropy are cornerstones of contemporary physics. They are both at the heart of thermodynamics, a set of connections that regulate the laws of the macroscopic world. Statistical mechanics bridges microscopic dynamics (Newtonian, quantum, relativistic, or any other) with the macroscopic behaviors we observe in nature. Boltzmann-Gibbs statistical mechanics constitutes an impressive illustration of how the logics of these micro-macro connections is structured. This theory is based on the BG expression  $S_{BG}$  for the entropy. Following along the lines of Einstein 1910 and of many others, it is our belief that the specific mathematical form that  $S_{BG}$  has must ultimately descend from microscopic dynamics. More precisely, this connection is known to rely on quick mixing

in phase space, ergodicity, and related nonlinear dynamical concepts. This is the statistical mechanics of short range interactions, short-range microscopic memory, ultimately leading to a simple, uniform occupancy of the allowed phase space (the “equal probability” hypothesis for the microcanonical ensemble). The question arises naturally: what happens when the interactions are long-ranged, when the microscopic memory is heavily non-markovian, when the geometry of occupation of phase space is complex, (multi)fractal for instance, or some other hierarchical structure? Can we still devise theoretical techniques similar to those of BG statistical mechanics, which would once again bridge with the laws of the macroscopic world? We believe that we can. A central point therefore is: what mathematical expression should or could we use for the entropy? How can we adequately generalize  $S_{BG}$ ?

The answer to such question obviously is far from trivial. Indeed, even in the framework of BG statistical mechanics, we do not know how to rigorously make the expression of  $S_{BG}$  descend from microscopic dynamics, plus possibly some other generic logical requests. In some sense, the best we know nowadays about such fundamental question lies at the level of the necessary and sufficient conditions proposed by Shannon, by Khinchin, and similar constructs. *But* they do not start from dynamics.

Since the direct connection between normal (ergodic) microscopic dynamics and  $S_{BG}$  is yet not totally clear, there is no surprise that the same happens for anomalous microscopic dynamics. This is the fundamental reason for which we have proceeded to generalize  $S_{BG}$  into  $S_q$  through a metaphoric path, as illustrated in Section II. The  $q$ -generalizations of the Shannon and the Khinchin theorems have already been established by Santos and by Abe respectively. A variety of arguments qualify the conditions under which  $S_q$  is unique (see [83] and references therein). Still, how to make  $S_q$  (or even  $S_{BG}$ , as we said before) descend from microscopic dynamics remains an open question. Consistently remains an open question what exactly have in common all the physical cases indicated in Section III, and what may exactly be the geometrical structure which reflects the dynamical occupancy of the phase space in such cases. We have conjectured that it may well be similar to a scale-free geometry like that of the Albert-Barabasi model, but this remains to be proved.

Another intriguing question is the connection with aging, a property that long-range-interacting Hamiltonians share with many glasses, spin-glasses, and other metastable systems intensively studied nowadays. Such long-range Hamiltonians satisfy, even in the longstanding aging phase (whose duration diverges with  $N$ ), a zeroth principle of thermodynamics, which concretely opens the path to thermodynamics. Furthermore, a sensible amount of connections with nonextensive statistical mechanics are already available in the literature. However, the exact dependence of  $q$  on  $(\alpha, d)$ , possibly on  $\alpha/d$ , is still to be unambiguously established. Many re-

searchers around the world are presently working on that fascinating problem.

Finally, we should emphasize that everything that we know today neatly points towards the scenario that nonextensive statistical mechanics satisfies the 0-th, first, second and third principle of thermodynamics (see, for instance, [95]). This is deeply interesting, since that implies that the basic laws of thermodynamics are stronger than the role reserved for them within Boltzmann-Gibbs statistical mechanics. Further analysis of the foundations and thermodynamical connections of statistical mechanics in general, and of nonextensive statistical mechanics in particular, are certainly fascinating and very welcome.

## VIII. ACKNOWLEDGMENTS

I have in more than one occasion referred in this paper to 1900 Planck’s contributions. This was in fact induced by delightful conversations I had with L. Tisza in 1995 (in H.E. Stanley’s office at the Boston University), and with L. Galgani in 2001 (during a Les Houches Winter School organized by T. Dauxois, S. Rufo, E. Arimondo and M. Wilkens). I warmly acknowledge here this double privilege.

I have also benefited from many invaluable discussions with M. Gell-Mann, as well as from many and interesting remarks from S. Abe, G.F.J. Ananos, C. Anteneodo, F. Baldovin, M. Baranger, E.P. Borges, E. Brigatti, M.O. Caceres, E.G.D. Cohen, E.M.F. Curado, L.G. Moyano, F.D. Nobre, A. Plastino, A.R. Plastino, A.K. Rajagopal, A. Rapisarda and A.M.C. Souza. Moreover, I am grateful to three anonymous Referees, whose helpful comments have resulted in the present, hopefully improved, version of the original manuscript.

Finally, without the intense and generous interest of H.L. Swinney and A.R. Bishop, and others, in organizing and supporting the Los Alamos National Laboratory *International Workshop on Anomalous Distributions, Non-linear Dynamics and Nonextensivity*, the present volume would have not existed.

- 
- [1] N.S. Krylov, *Nature* **153**, 709 (1944); N.S. Krylov, *Works on the Foundations of Statistical Physics*, translated by A.B. Migdal, Ya. G. Sinai and Yu. L. Zeeman, Princeton Series in Physics (Princeton University Press, Princeton, 1979).
  - [2] M. Planck, *Verhandlungen der Deutschen Physikalischen Gesellschaft* **2**, 202 and 237 (1900) (English translation: D. ter Haar, S. G. Brush, *Planck’s Original Papers in Quantum Physics* (Taylor and Francis, London, 1972)).
  - [3] S. Watanabe, *Knowing and Guessing* (Wiley, New York, 1969).

- [4] H. Barlow, *Vision. Res.* **30**, 1561 (1990).
- [5] D.H.E. Gross, *Microcanonical Thermodynamics: Phase Transitions in Small Systems*, Lecture Notes in Physics **66** (World Scientific, Singapore, 2001).
- [6] B. Lesche, *J. Stat. Phys.* **27**, 419 (1982).
- [7] C. Tsallis, in *Nonextensive Entropy – Interdisciplinary Applications*, eds. M. Gell-Mann and C. Tsallis (Oxford University Press, New York, 2004).
- [8] C. Tsallis, A.R. Plastino, and W.M. Zheng, *Chaos, Solitons and Fractals* **8**, 885 (1997).
- [9] F. Baldovin and A. Robledo, *Europhys. Lett.* **60**, 518 (2002), and *Phys. Rev. E* **66**, 8045104 (2002).
- [10] F. Baldovin and A. Robledo, *cond-mat/0304410* (2003). See also A. Robledo, *cond-mat/0307285*.
- [11] Y.S. Weinstein, S. Lloyd and C. Tsallis, *Phys. Rev. Lett.* **89**, 214101 (2002); Y.S. Weinstein, C. Tsallis and S. Lloyd, *On the emergence of nonextensivity at the edge of quantum chaos*, in *Decoherence and Entropy in Complex Systems*, ed. H.-T. Elze, *Lecture Notes in Physics* **633** (Springer, Heidelberg, 2004), page 385.
- [12] C. Tsallis, *J. Stat. Phys.* **52**, 479 (1988).
- [13] E.M.F. Curado and C. Tsallis, *J. Phys. A* **24**, L69 (1991) [Corrigenda: **24**, 3187 (1991) and 25, 1019 (1992)].
- [14] C. Tsallis, R.S. Mendes and A.R. Plastino, *Physica A* **261**, 534 (1998).
- [15] S.R.A. Salinas and C. Tsallis, eds., *Nonextensive Statistical Mechanics and Thermodynamics*, *Brazilian Journal of Physics* **29** (1999); S. Abe and Y. Okamoto, eds., *Nonextensive Statistical Mechanics and Its Applications*, *Series Lecture Notes in Physics* **560** (Springer, Berlin, 2001); G. Kaniadakis, M. Lissia and A. Rapisarda, eds., *Non Extensive Thermodynamics and Physical Applications*, *Physica A* **305** (Elsevier, Amsterdam, 2002); M. Gell-Mann and C. Tsallis, eds. *Nonextensive Entropy – Interdisciplinary Applications*, (Oxford University Press, New York, 2004). A regularly updated bibliography on the subject can be found at <http://tsallis.cat.cbpf.br/biblio.htm>.
- [16] M.L. Lyra and C. Tsallis, *Phys. Rev. Lett.* **80**, 53 (1998).
- [17] U.M.S. Costa, M.L. Lyra, A.R. Plastino and C. Tsallis, *Phys. Rev. E* **56**, 245 (1997).
- [18] F.A.B.F. de Moura, U. Tirnakli and M.L. Lyra, *Phys. Rev. E* **62**, 6361 (2000).
- [19] E.P. Borges, C. Tsallis, G.F.J. Ananos and P.M.C. Oliveira, *Phys. Rev. Lett.* **89**, 254103 (2002).
- [20] S. Abe, *Phys. Rev. E* **66**, 046134 (2002).
- [21] V. Latora, M. Baranger, A. Rapisarda, and C. Tsallis, *Phys. Lett. A*, **273** 97 (2000).
- [22] R. Albert and A.L. Barabasi, *Phys. Rev. Lett.* **85**, 5234 (2000); S. Abe and N. Suzuki, *Europhys. Lett.* **65**, 581 (2004).
- [23] C. Beck and F. Schlogl, *Thermodynamics of Chaotic Systems* (Cambridge University Press, Cambridge, 1993).
- [24] S. Abe and A.K. Rajagopal, *J. Phys. A* **33**, 8733 (2000). See also S. Abe and A.K. Rajagopal, *Europhys. Lett.* **52**, 610 (2000), *Phys. Lett. A* **272**, 341 (2000), *Europhys. Lett.* **55**, 6 (2001), and *Physica A* **295**, 172 (2001).
- [25] C. Budde, D. Prato and M. Re, *Phys. Lett. A* **283**, 309 (2001).
- [26] A.R. Plastino and A. Plastino, *Physica A* **222**, 347 (1995); C. Tsallis and D.J. Bukman, *Phys. Rev. E* **54**, R2197 (1996); E.M.F. Curado and F.D. Nobre, *Phys. Rev. E* **67**, 021107 (2003).
- [27] C. Anteneodo and C. Tsallis, *J. Math. Phys.* **44**, 5194 (2003).
- [28] S. Abe, *Phys. Rev. E* **68**, 031101 (2003).
- [29] A. Einstein, *Annalen der Physik* **33**, 1275 (1910) [“Usually  $W$  is put equal to the number of complexions... In order to calculate  $W$ , one needs a complete (molecular-mechanical) theory of the system under consideration. Therefore it is dubious whether the Boltzmann principle has any meaning without a complete molecular-mechanical theory or some other theory which describes the elementary processes.  $S = \frac{R}{N} \log W + \text{const.}$  seems without content, from a phenomenological point of view, without giving in addition such an *Elementartheorie*.” (Translation: Abraham Pais, *Subtle is the Lord ...* Oxford University Press, 1982)].
- [30] E.G.D. Cohen, *Physica A* **305**, 19 (2002); also his contribution in the present volume.
- [31] M. Baranger, *Physica A* **305**, 27 (2002).
- [32] U. Tirnakli, C. Tsallis and M.L. Lyra, *Phys. Rev. E* **65**, 036207 (2002).
- [33] U. Tirnakli, *Phys. Rev. E* **66**, 066212 (2002); E.P. Borges and U. Tirnakli, *cond-mat/0302616*.
- [34] F. Baldovin, E. Brigatti and C. Tsallis, *Phys. Lett. A* **320**, 254 (2004).
- [35] G.F.J. Ananos, F. Baldovin and C. Tsallis, *Anomalous sensitivity to initial conditions and entropy production in standard maps: A nonextensive approach*, preprint (2004).
- [36] G.A. Tsekouras, A. Provata and C. Tsallis, *Phys. Rev. E* **69**, 016120 (2004).
- [37] C. Anteneodo, private communication, and *cond-mat/0402248*.
- [38] B.M. Boghosian, P.J. Love, P.V. Coveney, I.V. Karlin, S. Succi and J. Yepez, *Phys. Rev. E* **68**, 025103(R) (2003); B.M. Boghosian, P. Love, J. Yepez and P.V. Coveney, *Galilean-invariant multi-speed entropic lattice Boltzmann models*, in the present volume.
- [39] M. Antoni and S. Ruffo, *Phys. Rev. E* **52**, 2361 (1995).
- [40] C. Anteneodo and C. Tsallis, *Phys. Rev. Lett.* **80**, 5313 (1998).
- [41] M.E. Fisher, *Arch. Rat. Mech. Anal.* **17**, 377 (1964), *J. Chem. Phys.* **42**, 3852 (1965) and *J. Math. Phys.* **6**, 1643 (1965); M.E. Fisher and D. Ruelle, *J. Math. Phys.* **7**, 260 (1966); M.E. Fisher and J.L. Lebowitz, *Commun. Math. Phys.* **19**, 251 (1970).
- [42] A. Campa, A. Giansanti, D. Moroni and C. Tsallis, *Phys. Lett. A* **286**, 251 (2001).
- [43] V. Latora, A. Rapisarda and C. Tsallis, *Phys. Rev. E* **64**, 056134 (2001).
- [44] M.A. Montemurro, F. Tamarit and C. Anteneodo, *Phys. Rev. E* **67**, 031106 (2003).
- [45] B.J.C. Cabral and C. Tsallis, *Phys. Rev. E* **66**, 065101(R) (2002).
- [46] C. Tsallis, *cond-mat/0304696* (2003), and *Phys. Rev. E* (1 March 2004), in press.
- [47] L.G. Moyano, F. Baldovin and C. Tsallis, *cond-mat/0305091*.

- [48] F.D. Nobre and C. Tsallis, Phys. Rev. E **68**, 036115 (2003).
- [49] E.P. Borges, C. Tsallis, A. Giansanti and D. Moroni, in *Tendencias da Fisica Estatistica no Brasil* [in Portuguese], (Editora Livraria da Fisica, Sao Paulo, 2003), page 84.
- [50] J.P.K. Doye, Phys. Rev. Lett. **88**, 238701 (2002).
- [51] E. Lutz, Phys. Rev. A **67**, 051402(R) (2003).Phys.
- [52] P. Erdos and J. Spencer, Comput. Math. Appl. **5**, 33 (1979).
- [53] N. Lemke and R.M.C. de Almeida, Physica A **325**, 396 (2003).
- [54] P.A. Alemany and D.H. Zanette, Phys. Rev. E **49**, R956 (1994); C. Tsallis, S.V.F. Levy, A.M.C. de Souza and R. Maynard, Phys. Rev. Lett. **75**, 3589 (1995) [Erratum: **77**, 5442 (1996)]; D. Prato and C. Tsallis, Phys. Rev. E **60**, 2398 (1999); S. Abe and A.K. Rajagopal, J. Phys. A **33**, 8723 (2000).
- [55] M. Bologna, C. Tsallis and P. Grigolini, Phys. Rev. E **62**, 2213 (2000).
- [56] C. Tsallis and E.K. Lenzi, in *Strange Kinetics*, eds. R. Hilfer et al, Chem. Phys. **284**, 341 (2002) [Erratum: **287**, 341 (2002)]; E.K. Lenzi, R.S. Mendes and C. Tsallis, Phys. Rev. E **67**, 031104 (2003).
- [57] M.O. Caceres, Phys. Rev. E **67**, 016102 (2003).
- [58] S. Gheorghiu, J.R. van Ommen and M.-O. Coppens, Phys. Rev. E **67**, 041305 (2003).
- [59] A. Baldassarri, U.M.B. Marconi and A. Puglisi, Europhys. Lett. **58**, 14 (2002).
- [60] C. Tsallis, G. Bemsiki and R.S. Mendes, Phys. Lett. A **257**, 93 (1999).
- [61] C. Tsallis, J.C. Anjos and E.P. Borges, Phys. Lett. A **310**, 372 (2003).
- [62] G. Wilk and Z. Wlodarczyk, Phys. Rev. Lett. **84**, 2770 (2000).
- [63] C. Beck, Phys. Rev. Lett. **87**, 180601 (2001).
- [64] C. Tsallis, in *New trends in magnetism, magnetic materials and their applications*, eds. J.L. Morán-López and J.M. Sanchez (Plenum Press, New York, 1994) page 451.
- [65] D. Prato, Phys. Lett. A **203**, 165 (1995).
- [66] E.K. Lenzi, R.S. Mendes and A.K. Rajagopal, Phys. Rev. E **59**, 1398 (1999).
- [67] C. Beck and E.G.D. Cohen, Physica A **321**, 267 (2003).
- [68] C. Tsallis and A.M.C. Souza, Phys. Rev. E **67**, 026106 (2003).
- [69] A.M.C. Souza and C. Tsallis, Phys. Lett. A **319**, 273 (2003).
- [70] A.M. Crawford, N. Mordant and E. Bodenschatz, physics/0212080.
- [71] T. Gotoh and R.H. Kraichnan, in this volume (2003).
- [72] A.M.C. Souza and C. Tsallis, in preparation.
- [73] G. Nicolis and D. Daems, Chaos **8**, 311 (1998).
- [74] B.P. Vollmayr-Lee and E. Luijten, Phys. Rev. E **63**, 031108 (2001).
- [75] C. Tsallis, cond-mat/0011022.
- [76] M.E. Fisher, Arch. Rat. Mech. Anal. **17**, 377 (1964); J. Chem. Phys. **42**, 3852 (1965); J. Math. Phys. **6**, 1643 (1965).
- [77] A. Pluchino, V. Latora and A. Rapisarda, to appear in *Anomalous distributions, nonlinear dynamics and nonextensivity*, eds. H.L. Swinney and C. Tsallis, Physica D (2004) [cond-mat/0303081].
- [78] Lj. Milanovic, H.A. Posch and W. Thirring, Phys. Rev. E **57**, 2763 (1998).
- [79] M. Antoni and A. Torcini, Phys. Rev. E **57**, R6233 (1998).
- [80] V. Latora, A. Rapisarda and S. Ruffo, Phys. Rev. Lett. **80**, 692 (1998); Physica D **131**, 38 (1999); Phys. Rev. Lett. **83**, 2104 (1999); Physica A **280**, 81 (2000).
- [81] M. Nauenberg, Phys. Rev. E **67**, 036114 (2003).
- [82] R. Luzzi, A.R. Vasconcellos and J. Galvao Ramos, Science **298**, 1171 (2002).
- [83] S. Abe, *Tsallis entropy: How unique?*, to appear in a special volume of Continuum Mechanics and Thermodynamics, ed. M. Sugiyama (2004), in press [cond-mat/0305087].
- [84] C. Tsallis, D. Prato and A.R. Plastino, *Nonextensive statistical mechanics: Some links with astronomical phenomena*, to appear in the Proceedings of the *XIth United Nations / European Space Agency Workshop on Basic Space Sciences*, Office for Outer Space Affairs / United Nations (Cordoba, 9-13 September 2002), eds. H. Haubold and M. Rabolli, special issue of Astrophysics and Space Science (Kluwer Academic Publishers, Dordrecht, 2004), in press [cond-mat/0301590].
- [85] C. Tsallis and E.P. Borges, in *Proc. 10th International Workshop on Multiparticle Production - Correlations and Fluctuations in QCD*, eds. N.G. Antoniou, F.K. Diakonos and C.N. Ktorides (World Scientific, Singapore, 2003), page 326.
- [86] S. Abe and A.K. Rajagopal, Science **300**, 249 (2003).
- [87] A. Plastino, Science **300**, 250 (2003).
- [88] V. Latora, A. Rapisarda and A. Robledo, Science **300**, 250 (2003).
- [89] D.H. Zanette and M.A. Montemurro, Phys. Rev. E **67**, 031105 (2003).
- [90] D.H. Zanette and M.A. Montemurro, Phys. Lett. A **316**, 184 (2003).
- [91] C. Tsallis, C. Anteneodo, L. Borland and R. Osorio, Physica A **324**, 89 (2003).
- [92] D.H. Zanette and M.A. Montemurro, cond-mat/0305070.
- [93] C. Tsallis, *Critique of the Zanette-Montemurro note concerning non thermodynamical applications of nonextensive statistics*, preprint (2003), unpublished.
- [94] F. Takens, in *Structures in dynamics - Finite dimensional deterministic studies*, eds. H.W. Broer, F. Dumortier, S.J. van Strien and F. Takens (North-Holland, Amsterdam, 1991), page 253. [“The values of  $p_i$  are determined by the following dogma: if the energy of the system in the  $i^{th}$  state is  $E_i$  and if the temperature of the system is  $T$  then:  $p_i = \exp\{-E_i/kT\}/Z(T)$ , where  $Z(T) = \sum_i \exp\{-E_i/kT\}$ , (this last constant is taken so that  $\sum_i p_i = 1$ ). This choice of  $p_i$  is called Gibbs distribution. We shall give no justification for this dogma; even a physicist like Ruelle disposes of this question as “deep and incompletely clarified.”]
- [95] S. Abe and A.K. Rajagopal, Phys. Rev. Lett. **91**, 120601 (2003).

ENERGY CONSTRAINT	BG	1988 VERSION (Tsallis)	1991 VERSION (Curado-Tsallis)	1998 VERSION (Tsallis-Mendes-Plastino)
PROPERTY	$\sum_i p_i E_i = U_{BG}$	$\sum_i p_i E_i = U_q$	$\sum_i p_i^q E_i = U_q$	$\frac{\sum_i p_i^q E_i}{\sum_i p_i^q} = U_q$
(i)	YES	YES	NO	YES
(ii)	YES	YES	NO	YES
(iii)	YES	NO	NO	YES
(iv)	YES	---	---	YES
(v)	YES	NO	YES	YES
(vi)	YES	NO	YES	YES
(vii)	YES	NO	YES	YES

FIG. 1. Successive versions for the canonical-ensemble energy-constraint within nonextensive statistical mechanics regarding properties (i-vii) indicated in Section III assuming fixed the Lagrange parameter  $\beta$ . In the 1988 paper [12] two possible constraints for the energy were indicated, namely  $\sum_i p_i E_i = \text{constant}$  and  $\sum_i p_i^q E_i = \text{constant}$ , but only the former was developed therein. In the 1991 paper [13], the second form for the constraint was developed, which enabled the connection with thermodynamics. In the 1998 paper [14], the constraint using escort distributions was adopted and developed, i.e.,  $\sum_i p_i^q E_i / \sum_i p_i^q = \text{constant}$ . This last form provides for the points (i-vii) the *same* answers as within BG statistical mechanics, and is presently believed to be its correct  $q$ -generalization.

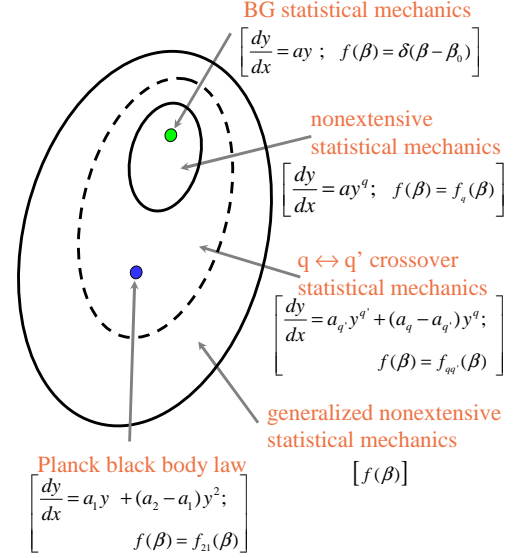


FIG. 2. Logical structure of the successive generalizations of *classical* BG statistical mechanics presented in this paper, with indication of their paradigmatic differential equations, and their  $f(\beta)$  superstatistical distributions;  $f_1(\beta) = f_{BG}(\beta)$ ;  $f_q(\beta) = \chi^2$ -distribution. The corresponding stationary state probability distributions and entropies are given by Eqs. (9) and (10), for BG statistical mechanics, by Eqs. (30) and (34), for nonextensive statistical mechanics, and by Eq. (86) and  $S_G$  given in [68,69], for generalized nonextensive statistical mechanics. The  $q \leftrightarrow q'$  crossover statistical mechanics includes as a particular case the  $q \leftrightarrow 1$  crossover one, which corresponds to Eq. (88). If we take into account in this equation the normalization factor, the  $\beta_2/\beta_1 \rightarrow \infty$  limit, and the photonic density of states, we recover Planck's law for the black-body radiation. See the text in Section V.C for clarification of this case.